

LANGUAGE

JOURNAL OF THE LINGUISTIC
SOCIETY OF AMERICA

EDITED BY BERNARD BLOCH

VOLUME 29

1953

Reprint with the permission of the Linguistic Society of America

KRAUS REPRINT CORPORATION

New York

1964

/

LINGUISTIC SOCIETY OF AMERICA

FOUNDED 1924 FOR THE ADVANCEMENT OF THE SCIENTIFIC STUDY
OF LANGUAGE; INCORPORATED 1940 IN THE DISTRICT OF COLUMBIA

OFFICERS AND COMMITTEES FOR 1953

President, BERNARD BLOCH, Yale University

Vice-President, J. ALEXANDER KERNS, New York University

Secretary and Treasurer, ARCHIBALD A. HILL, University of Virginia

Executive Committee, the preceding and

CARLETON T. HODGE, Foreign Service Institute (through 1953)

WERNER F. LEOPOLD, Northwestern University (through 1953)

JOSEPH H. GREENBERG, Columbia University (through 1954)

WILLIAM G. MOULTON, Cornell University (through 1954)

Committee on Publications

BERNARD BLOCH, Yale University, *Chairman and Editor*

RALPH L. WARD, Yale University (through 1953)

CHARLES F. HOCKETT, Cornell University (through 1954)

MARTIN JOOS, University of Wisconsin (through 1955)

Nominating Committee

YUEN REN CHAO, University of California (through 1953)

FRANKLIN EDGERTON, Yale University (through 1954)

HANS KURATH, University of Michigan (through 1955)

Standing Committee on Research

HENRY M. HOENIGSWALD, University of Pennsylvania (through 1953)

GEORGE S. LANE, University of North Carolina (through 1954)

HARRY HOIJER, University of California at Los Angeles (through 1955)

Delegate to the American Council of Learned Societies

J MILTON COWAN, Cornell University (through 1955)

Delegate to the American Association for the Advancement of Science

GEORGE L. TRAGER, Foreign Service Institute

Delegate to the Mexican Council for Indigenous Languages

NORMAN A. McQUOWN, University of Chicago

LANGUAGE is published quarterly by the LINGUISTIC SOCIETY OF AMERICA at the Waverly Press, Inc., Mt. Royal and Guilford Avenues, Baltimore 2, Md. Entered as Second Class Matter March 12, 1927, at the Postoffice at Baltimore, Md., under the Act of March 3, 1879; additional entry as Second Class Matter at the Postoffice at Charlottesville, Va.

Dues for Personal and Library Memberships in the Society are \$8.00 per calendar year. Of this sum \$5.00 is reserved for a subscription to LANGUAGE and its Supplements, and \$3.00 is reserved for Dues. Subscriptions are not sold without Membership.

Manuscripts for publication, exchange journals, and books for review or listing should be sent to the Editor of LANGUAGE (Bernard Bloch, Yale Graduate School, New Haven, Connecticut).

Applications for membership, library subscriptions, orders for current and back publications, etc., should be addressed to the Secretary of the Society (A. A. Hill, Box 1001, University Station, Charlottesville, Va.).

Copyright 1953 by LINGUISTIC SOCIETY OF AMERICA

Made in United States of America

CONTENTS OF VOLUME 29

I

Roland Grubb Kent (obituary by GEORGE S. LANE).....	1
JAAN PUHVEL: Indo-European negative composition.....	14
ALBERT MOREY STURTEVANT: Certain Old Norse secondary formations.....	26
E. COLIN CHERRY, MORRIS HALLE, and ROMAN JAKOBSON: Toward the logical description of languages in their phonemic aspect.....	34
YEHOShUA BAR-HILLEL: A quasi-arithmetical notation for syntactic description.....	47
MISCELLANEA:	
BERNARD BLOCH: Contrast.....	59
KEMP MALONE: Long and short in Icelandic phonemics.....	61
DWIGHT L. BOLINGER: Addenda to the comparison of inequality in Spanish.....	62
HERBERT H. PAPER: An Elamite etymology.....	66
REVIEWS:	
Shannon and Weaver: The mathematical theory of communication (CHARLES F. HOCKETT).....	69
Sturtevant: A comparative grammar of the Hittite language, revised edition (HOLGER PEDERSEN).....	93
Brunner: Die englische Sprache (HERBERT PENZL).....	96
Sauvageot: Esquisse de la langue hongroise (THOMAS A. SEBEOK).....	98
Polotsky: Notes on Gurage grammar (WOLF LESLAU).....	100
NOTES.....	104
PUBLICATIONS RECEIVED.....	107

II

ROBERT B. LEES: The basis of glottochronology.....	113
HORACE G. LUNT II: Old Church Slavonic <i>bedrno</i>	128
FREDERICK B. AGARD: Noun morphology in Romanian.....	134
SHERMAN M. KUHN and RANDOLPH QUIRK: Some recent interpretations of Old English digraph spellings.....	143
DAVID L. OLMSTED: Comparative notes on Yoruba and Lucumi.....	157
REVIEWS:	
Togebv: Structure immanente de la langue française (MURRAY FOWLER).....	165
Glinz: Die innere Form des Deutschen (WILLIAM G. MOULTON).....	175
Robins: Ancient and mediaeval grammatical theory in Europe (HENRY M. HOENIGSWALD).....	180
Slotty: Beiträge zur Etruskologie I (HENRY M. HOENIGSWALD).....	183
Leumann: Morphologische Neuerungen im altindischen Verbalssystem (LOUIS RENOU).....	186
Gonda: Remarques sur la place du verbe dans la phrase active et moyenne en langue sanscrite (LOUIS RENOU).....	187
Rodriguez Adrados: La dialectología griega como fuente para el estudio de las migraciones indoeuropeas en Grecia (CARL D. BUCK).....	188
Delebecque: Le cheval dans l'Iliade (JOSHUA WHATMOUGH).....	189
Devoto: Gli antichi Italici (JOSHUA WHATMOUGH).....	190
Pisani: Le lingue dell'Italia antica oltre il latino (JOSHUA WHATMOUGH).....	192
Snell: Der Aufbau der Sprache (WERNER WINTER).....	193
Kloss: Die Entwicklung neuer germanischer Kultursprachen von 1800 bis 1950 (WERNER WINTER).....	195
Hermodsson: Reflexive und intransitive Verba im älteren Westgermanischen (ALBERT M. STURTEVANT).....	197

Kurath and Kuhn, edd.: Middle English dictionary, Part E.1 (KEMP MALONE)....	204
Haguenauer: Morphologie du japonais moderne, Vol. 1 (JOSEPH K. YAMAGIWA)....	208
Malekebu, ed. Atkins: Unkhoswe waanyanja (DAVID L. OLMSTED).....	211
Panse, Kandler, Leischner: Klinische und sprachwissenschaftliche Untersuchungen zum Agrammatismus (DAVID L. OLMSTED).....	213
NOTES.....	215
PUBLICATIONS RECEIVED.....	222

III

Number dedicated to Franklin Edgerton

Dedication.....	227
JULES BLOCH: Prâkrit <i>cia</i> , latin <i>quidem</i>	229
LOUIS RENOU: Observations sur les composés nominaux du Rgveda.....	231
RULON WELLS: Secondary derivation from Sanskrit <i>i</i> -stems.....	237
E. ADELAIDE HAHN: Some Hittite-Sanskrit parallels.....	242
EMILE BENVENISTE: La flexion pronominale en hittite.....	255
ALBRECHT GOETZE: The theophorous elements of the Anatolian proper names from Cappadocia.....	263
GEORGE S. LANE: Imperfect and preterit in Tocharian.....	278
HENRY M. HOENIGSWALD: 'Pa, <i>śīdāe</i> , <i>śāśv</i> , and the semivowels.....	288
GEORGE M. BOLLING: Three puzzles in the language of the Iliad.....	293
JOSHUA WHATMOUGH: Epigraphica.....	297
GUSTAV MUST: The genitive singular of <i>o</i> -stems in Germanic.....	301
KONSTANTIN REICHARDT: The inscription on helmet B of Negau.....	306
ROBERT A. HALL JR.: The oaths of Strassburg: Phonemics and classification.....	317
MYLES DILLON: Semantic distribution in Gaelic dialects.....	322
GEORGE L. TRAGER: Russian declensional morphemes.....	326
MURRAY B. EMENEAU: Dravidian kinship terms.....	339
WILLIAM S. CORNYN: A Burmese Jātaka commentary.....	354
ISIDORE DYEN: Dempwolff's *R.....	359
GEORGE A. KENNEDY: Two tone patterns in Tangsic.....	367
KUN CHANG: On the tone system of the Miao-Yao languages.....	374
YUEN REN CHAO: Popular Chinese plant words.....	379

IV

W. F. TWADDELL: Stetson's model and the 'supra-segmental phonemes'.....	415
JAAN PUHVEL: Laryngeals and the Indo-European desiderative.....	454
ALBERT MOREY STURTEVANT: Further Old Norse secondary formations.....	457
ERIC H. LENNEBERG: Cognition and ethnolinguistics.....	463
REVIEWS	
Jakobson, Fant, Halle: Preliminaries to speech analysis (PAUL L. GARVIN).....	472
Pokorny: Indogermanisches etymologisches Wörterbuch, fasc. 7 (JOSHUA WHAT- MOUGH).....	481
Krause: Westtocharische Grammatik I (GEORGE S. LANE).....	483
Renou: Grammaire de la langue védique (FRANKLIN EDGERTON).....	497
Cimochowski: Le dialecte de Dushmani (ERIC P. HAMP).....	500
Ó Cúiv: Irish dialects and Irish-speaking districts (ERIC P. HAMP).....	512
Ó Cúiv: The Irish of West Muskerry; de Bhaldraithe: The Irish of Cois Fhairrge; Breatnach: The Irish of Ring (ERIC P. HAMP).....	517
Banta: Abweichende spät- und vulgärlateinische Perfektbildungen (ROBERT L. POLITZER).....	528

Maurer: A unidade da România ocidental (ROBERT A. HALL JR.)	530
Stimm: Studien zur Entwicklungsgeschichte des Frankoprovenzalischen (ROBERT A. HALL JR.)	532
Sganzini: Vocabulario dei dialetti della Svizzera Italiana I (ROBERT A. HALL JR.)	533
Caduff: Essai sur la phonétique du parler rhétoroman de la Vallée de Tavetsch (ROBERT A. HALL JR.)	534
Matluck: La pronunciación en el español del Valle de México (ISMAEL SILVA-FUENZALIDA)	535
Paiva Boléo: Os estudos de lingüística românica na Europa e na América desde 1939 a 1948 (LAWRENCE B. KIDDLE)	536
Manuppella: Os estudos de filologia portuguesa de 1930 a 1949 (LAWRENCE B. KIDDLE)	540
Brøndum-Nielsen: Studier og tydninger (W. P. LEHMANN)	541
Skard: Dativstudien (STEFÁN EINARSSON)	542
Myhre: Vokalsystemet i Iddemålet (EINAR HAUGEN)	544
Studies in honor of Albert Morey Sturtevant (ARCHIBALD A. HILL)	547
Kökeritz: Shakespeare's pronunciation (ARCHIBALD A. HILL)	549
Jones: The triumph of the English language (ARCHIBALD A. HILL)	561
Poldauf: On the history of some problems of English grammar before 1800 (PAUL L. GARVIN)	563
Cohen: The phonemes of English (GEORGE L. TRAGER)	564
Sundwall: Kleinasiatische Nachträge; Laroche: Recueil d'onomastique hittite (ALBRECHT GOETZE)	566
Kennedy: ZH guide (EDWARD H. SCHAFER)	568
Lekens: Dictionnaire ngbandi; Hulstaert: Dictionnaire français-lomongo (JOSEPH H. GREENBERG)	576
Dahl: Malgache et maanjan (ISIDORE DYEN)	577
Kroeber (ed.): Anthropology today (DAVID L. OLMSTED)	590
PUBLICATIONS RECEIVED	598

THEORY OF THE EARTH
AND ITS HISTORY
BY
J. D. DILLON
M.A., F.R.S.
OF THE UNIVERSITY OF CAMBRIDGE
AND
OF THE INSTITUTION OF CIVIL ENGINEERS
LONDON
LONGMANS, GREEN & CO., LTD.
1908

THEORY OF THE EARTH
AND ITS HISTORY
BY
J. D. DILLON
M.A., F.R.S.
OF THE UNIVERSITY OF CAMBRIDGE
AND
OF THE INSTITUTION OF CIVIL ENGINEERS
LONDON
LONGMANS, GREEN & CO., LTD.
1908

ROLAND GRUBB KENT

ROLAND GRUBB KENT died on June 27th, 1952, at the age of 75, after a long illness. He was a Signer of the Call that led to the founding of the Linguistic Society, and a Life Member of the Society since 1927. But he was far more than one of the Society's founders: from the beginning until 1940 he was its Secretary-Treasurer, devoting sixteen years of unselfish and often thankless labor to the service of linguistics in America. In 1941, at the end of this long period of stewardship, he was the Society's President.

It was Roland Kent also who presided at that first meeting on December 28th, 1924, in the American Museum of Natural History in New York—the meeting at which our Society was organized. Its success was due in large measure to his able chairmanship. What he said on being called to the chair¹ reflects his attitude toward his position and his duties throughout his career: he 'stated that he appreciated the honor and asked the help of all those present in an effort to conduct the meeting efficiently, without loss of valuable time.' As in the many meetings that he was to guide in later years, the important thing for him was the business at hand.

The newly founded Society was almost immediately plagued by scholarly and financial problems. That it survived its difficult infancy, that even in the bitter years of the depression it not only held its own but grew in membership and effectiveness, we owe above all to the unceasing efforts of three men: George Melville Bolling, the first Editor of *LANGUAGE*; Edward Howard Sturtevant, the creator of the Linguistic Institute; and Roland Grubb Kent, the man at the helm. One tangible result of the Society's growth as an agency for the fostering and publication of scientific work was the early establishment of the two series of supplements to *LANGUAGE*: the *Language Monographs* in 1925 and the *Language Dissertations* in 1927. As Leonard Bloomfield has said,² 'Only those who, before the year 1925, worked in almost complete isolation can appreciate the change that came about with the existence of the Linguistic Society.'

Roland Kent was born on February 24th, 1877, at Wilmington, Delaware, the son of Lindley and Anna (Grubb) Kent. His father, a business man in Wilmington, had seen service as a volunteer in the Civil War and had risen to the rank of major in the Union Army. To this man's desire that his son should have a classical education, Roland Kent owed his early bent for languages, especially for Latin and Greek.

Young Kent received his preparatory education at the Friends' School in Wilmington. At the age of 15 he entered Swarthmore College as a sophomore. Here he received his B.A. degree in 1895, a B.L. in 1896, and an M.A. in 1898. Immediately after leaving Swarthmore he went to Germany to continue his studies, as was the custom for classicists. The year 1899-1900 he spent at the

¹ Proceedings of the LSA, Lg. 1.9 (1925).

² Twenty-one years of the Linguistic Society, Lg. 22.2 (1946).

University of Berlin, the next year at the University of Munich. In Berlin he studied Greek epigraphy and Greek historians under Kirchhoff, and Attic orators under Wilamowitz-Möllendorf; in Munich he attended von Christ's 'philologisches Seminar', and studied archeology under Furtwängler, Roman historians under von Wölfflin, and Roman private life under von Müller. In 1902 he worked at the American School of Classical Studies in Athens. Later the same year he came to the University of Pennsylvania as Harrison Fellow in Classics, and received his doctorate there in 1903. In 1904 he was appointed Harrison Research Fellow in Classics. It was in this year that he married Miss Gertrude Freeman Hall, who survives him.

Kent's career as a teacher³ began in 1896 at Lower Marion High School in Ardmore, Pennsylvania. In 1904 he was appointed Instructor in Greek and Latin in the University of Pennsylvania, the institution where he was to spend the rest of his life. He became Assistant Professor of Comparative Philology in 1909, and Professor in 1916; in 1942 his title was changed, at his own request, to Professor of Indo-European Linguistics. He retired in 1947. In addition to his post at the University of Pennsylvania, he held several concomitant teaching positions: from 1910 to 1914 he was Lecturer in Sanskrit at Bryn Mawr College; in the summer of 1921 he was Visiting Professor at the Ohio State University; and on five occasions—1928 and '29 in New Haven, 1938 in Ann Arbor, 1941 and '42 in Chapel Hill—he was on the faculty of the Linguistic Institute.

The broad range of Kent's scholarly interests is reflected in the list of the learned bodies to which he belonged. He was a member not only of the Linguistic Society but also of the American Oriental Society, the American Philological Association, and the Classical Association of the Atlantic States, as well as of two local groups: the Oriental Club of Philadelphia (of which he was Secretary-Treasurer from 1909 to 1919, and twice President, 1919-20 and 1934-5) and the Classical Club of Philadelphia (of which he was President in 1909-10). His lifelong admiration of French scholarship and his particular friendship for France are attested by his membership in the Société de Linguistique de Paris, the Société des Études Latines, and the Association Guillaume Budé. In 1925 he lectured at the Sorbonne; the next year he was named Officier de l'Instruction publique, and in 1934 he was decorated Chevalier de la Légion d'Honneur. In the latter year, also, he was elected a Fellow of the American Academy of Arts and Sciences; and in 1941 he became a Vice-President of the Philological Society of Great Britain.

Kent is known to the world as a Classicist and an Iranist. His early training and the beginnings of his career reach back into that period, just before the turn of the century, when the study of Greek and Latin was still a *sine qua non* for the humanist, and when those two languages served as a point of departure for linguistics—in those days called simply philology, or at most comparative philology. Like many other classical linguists of his day—compare Sturtevant's concentration on Hittite—he chose a field of research outside the confines of Greek and Latin: Old Persian. The study of Old Persian was much to the fore in the first two decades of this century: witness Bartholomae's *Altiranisches*

³ See his own account, *Pedagogue's progress*, CW 42.82-7 (1948).

Wörterbuch, 1904; Tolman's *Ancient Persian lexicon and texts*, 1908; Weissbach's *Keilinschriften der Achämeniden*, 1911; Meillet's *Grammaire du vieux perse*, 1915; and Johnson's *Historical grammar*, 1917.⁴ There can be no doubt that it is for his work in Old Persian that Kent will be best remembered, at least outside the immediate field of Latin philology. It is most fortunate that the ripened fruit of Kent's Old Persian research has been distilled into a single work for succeeding generations of scholars: his *Old Persian grammar, texts, lexicon* (New Haven, 1950)—a great monument to his memory, perhaps the greatest. For this book alone, even if he had written nothing else, Indo-Europeanists and Iranists alike would be forever in his debt.

Despite his concern with Old Persian, however, Kent remained faithful to his first love, the Classics. In this field he will be remembered especially for his small popular volume, *Language and philology* (Boston, 1923); for his monograph *The textual criticism of inscriptions* (1926); and for his two famous textbooks, *The sounds of Latin* (1932, 3d ed. 1945) and *The forms of Latin* (1946). But it was as editor and translator of Varro *De lingua latina* (1938) that he displayed his scholarly acumen to best advantage; for this was a task to test his powers to their limit.⁵ Only two translations of Varro existed before Kent's English version: one French and one Italian. He showed, in this work, that he was not only a linguist but a great philologist as well, in the broadest sense of that term.

Roland Kent was a man of courage, integrity, and inflexible will. He worked and lived by the rules. He was fierce in attack and in defense; but whether he had his way or not, he carried out the will of the majority without gloating and without rancor. A rule might seem to him good or bad, but a rule was a rule: it was there to be followed until the majority saw fit to repeal it. Students and colleagues found him always ready to help them with their problems. Where his own knowledge allowed him to give advice, he was unstinting of his time and his counsel. With sham and pretense he had no patience. He was quick to detect it, and merciless in his attack. He did not suffer fools gladly.

Roland Kent was a faithful churchman, an Episcopalian of the Anglo-Catholic type. He was buried on June 30th from the Church of the Good Shepherd in Rosemont, Pennsylvania.

GEORGE S. LANE, *University of North Carolina*

THE FOLLOWING BIBLIOGRAPHY was prepared by Kent himself and has been edited by Bernard Bloch; it is assumed to be substantially complete. In addition to the works listed here, which he published as a private scholar, Kent published a number of items as Secretary-Treasurer of the Linguistic Society from 1925 to 1940. These include the Proceedings of the first seventeen annual meetings and the first three summer meetings of the Society (printed in *LANGUAGE*, Vols. 1-11, and in the *Bulletin*, Nos. 9-14); several unsigned obituaries of members (printed in the Notes and Personalia department of the journal); and a large

⁴ See in particular Kent's résumé, *The present state of Old Persian studies*, JAOS 56.208-21 (1936).

⁵ On the Herculean labor of editing and translating Varro, see Whatmough's review of Kent's work in CP 34.379-93.

number of small news items concerning the members' activities (also in the Notes and Personalia).

The following abbreviations are used in the bibliography:

AJA	American journal of archaeology
AJP	American journal of philology
AOS	American Oriental Society
AYB	American year book
CJ	Classical journal
CP	Classical philology
CR	Classical review
CW	Classical weekly
IF	Indogermanische Forschungen
IJ	Indogermanisches Jahrbuch
JAOS	Journal of the American Oriental Society
JNES	Journal of Near Eastern studies
Lg.	Language
LSA	Linguistic Society of America
MLJ	Modern language journal
PAPA	Proceedings of the American Philological Association
PAPS	Proceedings of the American Philosophical Society
PG	Pennsylvania gazette (publication of UP)
PQ	Philological quarterly
TAPA	Transactions of the American Philological Association
UP	University of Pennsylvania

- 1899 The ideal school from a teacher's point of view, Montgomery County (Pa.) Teachers' Institute 1898.186-90.
- 1902 A visit to Sparta, Whittier miscellany (Friends' School, Wilmington, Del.), 29 March, pp. 15-7.
- 1903 Chalois and Euripus from the mainland, The critic 33.12-3 (Hopkins Grammar School, New Haven, Conn.).
On Albinovanus Peto vv. 1-7 apud Sen. Suas. I 15, CR 17.311-2.
- 1904 A history of Thessaly from the earliest historical times to the accession of Philip V of Macedonia (doctoral dissertation, privately printed in part), pp. viii, 27; Lancaster, Pa.
- 1905 The date of Aristophanes' birth, CR 19.311-2.
The City Gates of Demetrias, AJA 9.166-9.
- 1906 When did Aristophanes die? CR 20.153-5. Same in abstract form, PAPA 36.xlvii-xlviii.
- 1907 The time element in the Greek drama, TAPA 37.39-52.
- 1909 Hysteron proteron in the Aeneid I-VI, CW 3.74-8.
- 1910 Note on *primus*, Aeneid 1.1, CW 3.150.
- 1911 Indo-Germanic philology, AYB 1910.759-60.
Lucilius on *EI* and *I*, AJP 32.272-93.
The etymology of Syriac *dastabīrā*, JAOS 31.359-64.
The etymology of Latin *mīles*, TAPA 41.5-9.
Note on *haec ubi agrestem pepulere*, Horace, Sat. II, 6.97-8, PAPA 41.xlv-xlvi.
- 1912 Indo-Germanic philology, AYB 1911.778-80.
Zu den orthographischen Regeln des Lucilius, Glotta 4.299-302.

- Latin *mille* and certain other numerals, TAPA 42.69-89.
 Note on *malis ridentem alienis*, Horace, Sat. II, 3, 72, PAPA 42.xxx-xxxii.
 Review of H. E. Burton, *A Latin grammar*, CW 5.154-6, 162-4; reply to H. F. Fuller's rejoinder (CW 6.37-8), CW 6.38-9.
- 1913** [With I. F. Hall] *Stories from the Far East*, 153 pp.; New York: Charles E. Merrill Co.
 [Editor] *The Oriental Club of Philadelphia: Record of 25 years*, pp. 34.
 The Oscan slingshot of Saepinum, IF 196-202.
 Indo-European philology, AYB 1912.769-71.
 The Vedic path of the gods and the Roman pontifex, CP 8.317-26.
 Again Lucilius on *EI* and *I*, AJP 34.315-21.
 Classical parallels to a Sanskrit proverb, JAOS 33.214-6.
 The chronology of certain Indo-Iranian sound-changes, JAOS 33.259-62.
 Dissimilative writings for *II* and *III* in Latin, TAPA 43.35-56.
 Some purpose clauses: Xen. Anab. I.6.6., Verg. Aen. II.667, CJ 9.35-6.
 La provenance de quatre mots géorgiens, Revue de linguistique 46.267-70.
 Some tense sequences in Caesar, *De bello gallico*, CW 7.77-8. Correction, CW 7.88.
- 1914** Lateinisches *povero* 'puero', IF 33.169-71.
 Indo-European philology, AYB 1913.905-7.
EI-readings in the mss. of Plautus vs. mathematics, CP 9.199-200.
 The etymological meaning of *pomerium*, TAPA 44.9-24.
 The military tactics of Caesar and of today, CW 8.69-70.
 Review of J. Marouzeau, *Notes sur la fixation du latin classique*, CP 9.224.
 Review of M. A. Stewart, *A study on Latin abstract substantives*, CW 7.206-8.
 Review of J. C. Rolfe, *Suetonius with an English translation*, Vol. I, Old Penn 12.1109.
 Review of same, Vol. II, Old Penn 13.19.
 Review of C. Juret, *Dominance et résistance dans la phonétique latine*, CW 8.22-3.
- 1915** Note on Atharva-Veda XX.127.10, JAOS 34.310-2.
 Persönliche Nachrichten, IJ 2.216.
 Philologenversammlungen in den Vereinigten Staaten von Amerika 1913 bis Juni 1914, IJ 2.217-35.
 Note on Iliad 2.260, CJ 10.172-4.
 Indo-European philology, AYB 1914.769-70.
 The plutei to the front, CW 8.136.
 An offering to the Manes, CW 8.144.
 The Latin Department's collection of antiquities: II. Roman dentistry, Old Penn 13.964-7.
 The passing of the sequence of tenses, CW 9.2-7, 10-3.
 A modern epistolary past, CW 9.47.
 Review of L. Alexander, *The kings of Lydia*, CW 8.207-8.
 Review of E. D. Cressman, *The semantics of -mentum, -bulum, and -culum*, CW 9.72.
- 1916** Folk-tales of India, Lectures by members of the faculty (UP) 1915-6, 3.237-61.
 Indo-European philology, AYB 1915.765-6.
 Philological meetings in the United States of America to June 1915, IJ 3.203-13.
 Cunaxa and Ctesiphon, CW 9.160.
 More war notes, CW 9.184.
 The sequence of tenses: Rejoinder, CW 9.195-6, 198.
 The Latin Department's collection of antiquities: IV. The water supply of ancient Rome, Old Penn 14.945-7.
 The Latin Department's collection of antiquities: V. Roman lamps, Old Penn 14.1134-7.
 The turre ambulatoria and the perambulating tank, CW 10.48.
 [With E. H. Sturtevant] Elision and hiatus in Latin prose and verse, TAPA 46.129-66.
 Parallels, CW 10.71-2.

- Review of H. Güntert, *Über Reimwortbildungen im Arischen und Altgriechischen*, CW 9.166-7.
- Review of A. Ernout, *Morphologie historique du latin*, CP 11.246-8.
- 1917 Roman and modern military highways, CW 10.88.
- Indo-European philology, AYB 1916.771-2.
- Latin and Greek in the newspapers, Old Penn 15.388-90.
- Path (of the gods or of the soul), *Hastings' encyclopaedia of religion and ethics* 9.671-2.
- Review of E. W. Nichols, *The semantic variability and semantic equivalents of -oso- and -lento-*, CP 12.113-5.
- Review of W. M. Lindsay, *A short historical Latin grammar*², CW 10.188-90.
- 1918 Studies in the Old Persian inscriptions, JAOS 35.321-52.
- Zoroastrianism, *Religions of the past and present* 183-210 (Philadelphia: J. B. Lippincott Co.).
- Then and now, Alumni review (UP) 20.271-80.
- Indo-European philology, AYB 1917.737-8.
- [Editor] *Thirty years of Oriental studies*: Issued in commemoration of thirty years of activity of the Oriental Club of Philadelphia, 84 pp.; Philadelphia.
- Review of E. H. Sturtevant, *Linguistic change*, AJP 39.316-21.
- Review of E. L. Johnson, *Historical grammar of the ancient Persian language*, AJP 39.322-3.
- Review of A. H. M. Stonecipher, *Graeco-Persian names*, AJP 39.323.
- 1919 Indo-European philology, AYB 1918.784-5.
- Horace on the high seas, CW 13.41-5.
- Cattle-tending and agriculture in the Avesta, JAOS 39.329-38.
- 1920 Indo-European philology, AYB 1919.793-5.
- The need of a publication fund, PG 18.768.
- University duty of original research, PG 18.795-6.
- Unpublished research, PG 18.824.
- Some big timbers of antiquity, American lumberman 2348.63.
- English word formation and etymology, Pennsylvania State Educational Association 213-7.
- The Latin language in the fourth century, TAPA 50.91-100.
- Studies in the Iguvine Tables, CP 15.354-69.
- The edict of Diocletian fixing maximum prices, UP law Review 69.35-47.
- The textual criticism of inscriptions, JAOS 40.289-99.
- 1921 Some features of word composition in Lithuanian, Slavonic studies 2.3.3-10.
- Addendum on a difficult Old Persian passage, JAOS 41.74-5.
- Horace, Odes 4.6.1-28: Another study in punctuation, CW 14.144.
- Deciphers Roger Bacon manuscript, PG 19.851-3.
- The alleged conflict of the accents in Latin verse, TAPA 51.19-29.
- Review of E. H. Sturtevant, *The pronunciation of Greek and Latin*, AJP 42.183-5.
- Review of *Edict of Diocletian establishing a maximum schedule of prices for commodities and services throughout the Roman Empire 301 A.D.*, published by the Union Trust Company of Providence, R. I., CW 15.15-6.
- 1922 Latin as the international auxiliary language, CJ 18.38-44.
- Review of A. F. Braunlich, *The indicative indirect question in Latin*, CW 15.87-8.
- 1923 *The baffled Hercules from Sparta*, pp. 20; Lancaster, Pa.: Lancaster Press, 1923.
- [Reprint of a badly printed article in Proceedings of the Numismatic and Antiquarian Society of Philadelphia 1919-20-21, 29.85-104.]
- Latin as the international auxiliary language*, pp. 31; Princeton, N. J.: American Classical League.
- Language and philology* ('Our debt to Greece and Rome', Vol. 22), pp. 174; Boston: Marshall Jones Co.
- A modern Ruso, CW 18.128.
- Variety and monotony in Plautine plots, PQ 2.164-72. Addendum to same, PQ 2.315.
- The educated Roman and his accent, TAPA 53.63-72.

- Addendum on Catullus' *passer*, AJP 44.353-4.
 Review of H. Güntert, *Kalypso*, PQ 2.73-4.
 Review of H. Güntert, *Von der Sprache der Götter und Geister*, PQ 2.74-5.
 Review of O. Jespersen, *Language*, MLJ 7.314-6.
- 1924** The Association Guillaume Budé, CW 18.17-8.
 Words and their vagaries, Juntó (student publication, UP) 2.3.3-5.
 The scientist and an international language, PAPS 63.162-70.
 The scientist and an international language, Science 59.554-5, 60.202-3.
 Likes and dislikes in elision, and the Vergilian Appendix, TAPA 54.86-97.
 Stamps and Latin, Collectors Club philatelist 3.150-6.
 Review of F. Edgerton, *The Panchatantra reconstructed*, PG 23.58-9.
 Review of F. Edgerton, *The Panchatantra reconstructed*, New York Evening Post, 25 October, Literary review; Philadelphia Public Ledger, 26 October, Literary review.
- 1925** The LSA, PG 23.294.
 The LSA, Hispania 8.132-3.
 The LSA, Quarterly journal of speech education 11.199-200.
 [With C. R. J. Scott] Latin cross word puzzles, Language notes supplement, Service Bureau for classical teachers (Teachers College, New York) No. 6, i-1025.
 On the reading of Latin verse, CW 18.144.
 The Oscan curse of Vibia, CP 20.243-67.
 [With J. R. Ware] The Old Persian cuneiform inscriptions of Artaxerxes II and Artaxerxes III, TAPA 55.52-61.
 Final -ae in Latin case-forms, Lg. 1.103-6.
 On certain personal pronouns in the Indo-Iranian dialects, *Indo-Iranian studies in honour of Sanjana* 291-3 [found only in copies with the supplement].
 La Société de linguistique d'Amérique, Bulletin de l'Association Guillaume Budé, No. 9, 49-54.
Caliga 'chaussure de soldat', Bulletin de la Société de linguistique de Paris 26.110-2.
 L'accentuation latine: Problèmes et solutions, Revue des études latines 3.204-14.
 La lecture à haute voix des vers latins, Revue internationale de l'enseignement 79.321-35.
 The classical movement in France, CJ 20.356-8.
 Review of L. A. Waddell, *The Phoenician origin of Britons, Scots, and Anglo-Saxons*, JAOS 45.172-4.
 Review of F. Stählin, *Das hellenische Thessalien*, CP 20.192.
- 1926** *The textual criticism of inscriptions* (Language monograph No. 2), 76 pp.; Philadelphia: LSA.
 Correspondence: The decennial bibliography, AJP 47.104-5.
 The LSA, Hispania 9.501.
 Indo-European philology, AYB 1925.1019-21.
 International auxiliary language, AYB 1925.1021-4.
 [With Edith Hall Dohan] New inscriptions from Cyprus, AJA Ser. 2, 30.249-58.
 On some animal names in Italic, Lg. 2.184-80.
 The inscription of Duenos, Lg. 2.207-22.
 [With E. H. Sturtevant] *Survey of linguistic studies: Opportunities for advanced work in the United States* (Bulletin No. 1), pp. 16; Baltimore: LSA.
 Review of Leumann-Stolz, *Lateinische Grammatik: Laut- und Formenlehre*⁴, CJ 22.69-74.
 Review of E. Littmann, *Morgenländische Wörter im Deutschen*³, JAOS 46.248-9.
 Review of C. Autran, *Introduction à l'étude critique du nom propre grec*, JAOS 46.249-52.
 Review of A. Meillet, *Trois conférences sur les Gâthas de l'Avesta*, JAOS 46.273-4.
- 1927** [Editor] *Newbold memorial meeting*, pp. 30; Baltimore: Waverly Press. Also: General magazine of the UP 29.185-203.
 Indo-European philology, AYB 1926.1009-11.

- International auxiliary language, AYB 1926.1018-9.
 Summary of Glotta, Vol. 15, AJP 48.86-9, 181-2; of Glotta, Vol. 16, AJP 48.376-9.
 The *-tt-* in Latin *quattuor*, Lg. 3.12-4.
 Word contamination in the Italic dialects, TAPA 57.51-9.
 Note on Yasna 34.11, JAOS 47.267-8.
 Oscan *deketasio-*, Lg. 3.187.
 Review of A. Ernout, *Morphologie historique du latin*², AJP 48.92-3.
 Review of E. Goldmann, *Die Duenos-Inschrift*, AJP 48.93-6.
 Review of C. Callet, *Le mystère du langage*, Lg. 3.150-1.
 Review of J. Marouzeau, *Dix années de bibliographie classique 1914-24*, Vol. 1, AJP 48.182-4.
 Review of Ernout-Riemann, *Syntaxe latine*², AJP 48.187-9.
 Review of A. Carnoy, *La science du mot*, Lg. 3.199-202.
 Review of P. Chantraine, *Histoire du parfait grec*, Lg. 3.202-4.
 Review of W. W. Coxe, *The influence of Latin on the spelling of English words*, CW 21.15-6.
 Review of F. Müller, *Altitalisches Wörterbuch*, Lg. 3.262-72.
 Review of A. Klingenheben, *Die Laute des Ful*, Lg. 3.273-4.
 1928 [Editor] *The cipher of Roger Bacon*, by William Romaine Newbold, edited with foreword and notes, pp. xxxii, 224; Philadelphia: UP Press.
 A linguistic institute, News letter of the Philadelphia Teachers Association 24.27-9.
 Three notes on the Gathas of the Avesta, Lg. 4.106-8.
 Summary of Glotta, Vol. 16, AJP 49.203-5.
 Lachmann's law of vowel lengthening, Lg. 4.181-90.
 [With E. H. Sturtevant] Linguistic science and classical philology: The Linguistic Institute, session of 1928, CW 22.9-13.
 A syntactical shift in an Avestan passage, JAOS 48.277-9.
 The conquests of the Latin language, CJ 24.191-212.
 Review of de Vreese, *Petron 39 und die Astrologie*, CW 21.119.
 Review of G. Atherton, *The immortal marriage*, CW 21.165-7.
 Review of H. F. Standerwick, *Latin roots*, CW 21.181-2.
 Review of *The English-Latin debt* [anonymous], CW 21.182-3.
 Review of G. Lodge, *Classical origins of scientific terms*, CW 21.206-8.
 Review of *Festschrift Meinhof*, Lg. 4.128-30.
 Review of Z. Schapira, *Die Bibel als Ariadnefaden im Labyrinth der Sprachen*, Lg. 4.150.
 Review of Indogermanisches Jahrbuch XI, Lg. 4.204.
 Review of Portucale I, Lg. 4.210.
 Review of A. S. Pedreira, *De los nombres de Puerto Rico*, Lg. 4.210-11.
 Review of J. Morris, *The name of Oglethorpe*, Lg. 4.211-2.
 Review of W. A. Read, *Louisiana place-names of Indian origin*, Lg. 4.212-3.
 Review of W. Gregory, *Union list of serials*, Lg. 4.213.
 Review of J. D. C. Pavry, *The Zoroastrian doctrine of a future life*, JAOS 48.285-6.
 Review of A. Meillet, *Les langues dans l'Europe nouvelle*², Lg. 4.289-91.
 Review of S. Ichikawa, *English influence on Japanese*, Lg. 4.292.
 Review of K. Stegmann von Pritzwald, *Sprache und Persönlichkeit*, Lg. 4.292-3.
 1929 The cookery inscription from Praeneste, Lg. 5.18-22.
 Summary of Glotta, Vols. 17-18, AJP 50.86-9, 295-6, 390-3.
 The labial nasal before stops in Primitive Indo-European, *Donum natalicium Schrijnen* 342-6; Nijmegen and Utrecht.
 Notice of the Linguistic Institute, second session, Quarterly journal of speech education 15.461-2.
 Review of A. Meillet, *Esquisse d'une histoire de la langue latine*, Lg. 5.38-44.
 Review of K. Jaberg and J. Jud, *Sprach- und Sachatlas Italiens und der Südschweiz*, Lg. 5.45.

- Review of E. Lokotsch, *Etymologisches Wörterbuch der europäischen (germanischen, romanischen, und slavischen) Wörter orientalischen Ursprungs*, JAOS 49.93-4.
- Review of H. Bauche, *La langue populaire*², Lg. 5.123-5.
- Review of E. Pradez, *Dictionnaire des gallicismes*, Lg. 5.126-7.
- Review of A. Dauzat, *Les noms de personnes*², Lg. 5.127.
- Review of E. Goldmann, *Beiträge zur Lehre vom indogermanischen Charakter der etruskischen Sprache I*, AJP 50.215-6.
- Review of A. Senn, *Kleine Litauische Sprachlehre*, Lg. 5.194.
- Review of F. Specht, *Szyruids Punktay Sakimu*, Lg. 5.195.
- Review of J. Schrijnen, *De latijnsche genitief singularis van de o-stammen*, Lg. 5.196-7.
- Review of A. V. W. Jackson, *Zoroastrian studies*, JAOS 49.286-9.
- Review of H. Hirt, *Indogermanische Grammatik V: Der Akzent*, Lg. 5.262-5.
- Review of A. C. Juret, *La phonétique latine*, Lg. 5.290-3.
- 1930 On long *i* in Latin terminations, Lg. 6.304-18.
- Summary of Glotta, Vols. 18-19, AJP 51.191-2, 386-8.
- Review of G. K. Zipf, *Relative frequency as a determinant of phonetic change*, Lg. 6.85-6.
- Review of P. Crouzat and A. Fournier, *Du latin à l'anglais*, Lg. 6.93-4.
- Review of E. Goldmann, *Beiträge zur Lehre vom indogermanischen Charakter der etruskischen Sprache II*, AJP 51.308-9.
- Review of *Studies in honor of Hermann Collitz*, Lg. 6.319-21.
- Review of *Festschrift der 57. Versammlung deutscher Philologen und Schulmänner in Salzburg 1929 gewidmet*, Lg. 6.321-2.
- Review of A. Thumb, *Handbuch des Sanskrit mit Texte und Glossar*² I, revised by H. Hirt, Lg. 6.328-9.
- 1931 Summary of Glotta, Vol. 19, AJP 52.281-3.
- The alleged strength of the initial syllable, Lg. 7.179-89.
- 'No Trespass' in Latin linguistics, *Classical studies in honor of John C. Rolfe* 143-61; Philadelphia: UP Press.
- The recently published Old Persian inscriptions, JAOS 51.189-240. Reprinted as a Special Publication of the LSA.
- Review of A. Walde, *Lateinisches etymologisches Wörterbuch*², revised by J. B. Hofmann, fasc. 1-2, AJP 52.81-5; fasc. 3-4, AJP 52.366-8.
- Review of T. G. Tucker, *Notes on Indo-European etymologies*, Lg. 7.64-5.
- Review of E. Richter, *Über die Reihenfolge der Organeinstellung beim Sprechen*, Lg. 7.150-1.
- Review of A. Meillet, *Aperçu d'une histoire de la langue grecque*², Lg. 7.154-6.
- Review of J. Karst, *Les Ligures comme substratum ethnique dans l'Europe illyrienne et ouralo-hyperboréenne*, AJA 35.483-4.
- Review of M. Niedermann, *Précis de phonétique historique du latin*, Lg. 7.284.
- Review of J. Friedrich, *Hethitisch und 'kleinasiatische' Sprachen*, and E. Fiesel, *Etruskisch*, Lg. 7.286-7.
- 1932 *The sounds of Latin* (Language monograph No. 12), pp. 216; Baltimore: LSA.
- The development of the Indo-European dental groups, Lg. 8.18-26.
- Summary of Glotta, Vol. 20, AJP 53.176-8.
- Summaries of articles on the Forum stele, *Philologus* 86.460-91 (Stroux) and *Klio* 1932.1-90 (Graffunder-Laufer, Goldmann), AJA 36.548.
- Review of W. W. Goodwin, *Greek grammar*, revised by C. B. Gulick, CJ 27.381-4.
- Review of A. Meillet, *Esquisse d'une histoire de la langue latine*², Lg. 8.56-7.
- Review of F. M. Stawell, *A clue to the Cretan scripts*, AJA 36.217-8.
- Review of A. Ernout and A. Meillet, *Dictionnaire étymologique de la langue latine*, Lg. 8.152-65.
- Review of B. D. Meritt, *Corinth: Results of excavations conducted by the American School of Classical Studies at Athens* (Vol. 8, part 1, Greek inscriptions 1896-1927), AJA 36.369.

- 1933** The record of Darius's palace at Susa, *JAOS* 53.1-23; addendum 53.166.
 A new inscription of Xerxes, *Lg.* 9.35-46.
 Another inscription of Xerxes, *Lg.* 9.229-31.
 Summary of Glotta, Vol. 21, *AJP* 54.75-8, 384-6.
 Review of A. Meillet, *Grammaire du vieux-perse*², revised by E. Benveniste, *Lg.* 9.92-6.
 Review of B. C. F. Atkinson, *The Greek language*, *CJ* 28.458-60.
 Review of Association Guillaume Budé, *Congrès de Nîmes ... Actes du Congrès*, *CJ* 28.620-2.
 Review of A. von Blumenthal, *Die iguvinischen Tafeln*, *Lg.* 9.214-8.
 Review of A. Walde, *Lateinisches etymologisches Wörterbuch*², revised by J. B. Hofmann, fasc. 5-6, *AJP* 54.297-9.
 Review of C. D. Buck, *Comparative grammar of Greek and Latin*, *CJ* 29.63-6.
 Review of W. H. Buckler and D. M. Robinson, *Greek and Latin inscriptions* (Sardis, Vol. 8), *AJA* 37.516-7.
- 1934** The name Ahuramazda, *Oriental studies in honour of C. E. Pavry* 200-8; London: Oxford University Press.
 Latin *altus* in a partitive meaning, *Miscelânea científica e literária dedicada ao Doutor J. Leite de Vasconcellos* 1.110-5; Coimbra: Imprensa da Universidade.
 More Old Persian inscriptions, *JAOS* 54.34-52.
 William Romaine Newbold (article), *Dictionary of American biography* 13.448-9.
 Summary of Glotta, Vol. 22, *AJP* 55.275-7.
 Review of L. Bloomfield, *Language*, *Lg.* 10.40-8.
 Review of C. Arbenz, *Die Adjektive auf -ιος*, *Lg.* 10.52.
 Review of P. Chantraine, *La formation des noms en grec ancien*, *Lg.* 10.207-9.
 Review of A. Walde, *Lateinisches etymologisches Wörterbuch*², revised by J. B. Hofmann, fasc. 7, *AJP* 55.290-1.
 Review of A. C. Juret, *Système de la syntaxe latine*², *Lg.* 10.297-9.
 Review of A. Meillet, *Le slave commun*², revised by A. Vaillant, *Lg.* 10.300-1.
- 1935** Die oskisch-umbrische Personalendung -ns, *IF* 53.41-4.
 Linguistic science and the orientalist (presidential address, meeting of the AOS, Ann Arbor, 24 April), *JAOS* 55.115-37.
 Again the initial syllable, *Lg.* 11.151-2.
 The Firdausi celebration, *UP library chronicle* 3.19-20. Also *PG* 33.441-4.
 Summary of Glotta, Vol. 22, *AJP* 56.257-8.
 The etymology of Greek *thanatos* and its kin, *Lg.* 11.207-10.
 Review of A. Debrunner, *Nachklassisches Griechisch*, *Lg.* 11.42.
 Review of J. F. K. Dirichs, *Die urlateinische Reklamestrophe auf dem sogenannten Dresselschen Drillingsgefäß des sabinischen Töpfers Dufnos* (bisher *Duenos*), *Lg.* 11.43-4.
 Review of L. Tesnière, *La lutte des langues en Prusse orientale*, *Lg.* 11.49.
 Review of H. Hirt, *Indogermanische Grammatik VI: Syntax I*, *Lg.* 11.154-60.
 Review of A. Meillet, *Introduction à l'étude comparative des langues indo-européennes*¹, *Lg.* 11.160-1.
 Review of *Conférences de l'Institut de linguistique de l'Université de Paris*, Année 1933, *Lg.* 11.161-2.
 Review of E. Benveniste and L. Renou, *Vṛtra et Vṛθragna*, *Lg.* 11.162-3.
 Review of O. Skutsch, *Prosodische und metrische Gesetze der Iambenkürzung*, *Lg.* 11.163.
 Review of A. W. Read, *Lexical evidence from folk epigraphy in western North America*, *Lg.* 11.166.
- 1936** The present status of Old Persian studies, *JAOS* 56.208-25. Reprinted: AOS reprint series, No. 7.
 Assimilation and dissimilation, *Lg.* 12.245-58.
 Review of *Atti del III Congresso internazionale dei linguisti*, *Lg.* 12.64.

- Review of E. Schwyzer, *Griechische Grammatic*, fasc. 1, Lg. 12.65-9.
 Review of A. Prévot, *L'aoriste grec en -ον*, Lg. 12.69-71.
 Review of R. Nehrbass, *Sprache und Stil der Iamata von Epidauros*, Lg. 12.71.
 Review of *Thesaurus linguae latinae epigraphicae* 2.1, Lg. 12.72.
 Review of A. C. Baugh, *A history of the English language*, Lg. 12.72-5.
 Review of G. Margouliès, *Petit précis de grammaire chinoise écrite*, Lg. 12.79.
 Review of G. Dumézil, *Textes populaires inguś*, Lg. 12.80.
 Review of M. Honnorat, *La langue gauloise resuscitée*, Lg. 12.81.
 Review of F. Butavand, *L'énigme lydienne*, Lg. 12.81.
 Review of A. Walde, *Lateinisches etymologisches Wörterbuch*³, revised by J. B. Hofmann, fasc. 8, AJP 57.225.
 Review of J. W. Poultny, *The syntax of the genitive case in Aristophanes*, Lg. 12.302-3.
 Review of N. Fukushima, *On the designation-problem of the so-called Tocharian language*, Lg. 12.303.
 Review of E. Benveniste, *Les infinitifs avestiques*, Lg. 12.303-4.
 Review of O. W. Heick, *The ab urbe condita construction in Latin*, Lg. 12.305-6.
 Review of N. Bachtin, *Introduction to the study of modern Greek*, Lg. 12.306-7.
 Review of I. C. Ward, *An introduction to the Ibo language*, Lg. 12.307-8.
 Review of C. R. Noyes, *Etymology of early legal terminology*, Lg. 12.308-9.
1937 On the text of Varro *De lingua latina*, TAPA 67.64-82.
 The Daiva-inscription of Xerxes, Lg. 13.292-305.
 Latin *tepīdus*, Spanish-Portuguese *tibio*, Lg. 13.145-6.
 Review of *Conférences de l'Institut de linguistique de l'Université de Paris*, Année 1934, Lg. 13.76-7; Année 1935, Lg. 13.318.
 Review of F. W. König, *Die Stele von Xanthos*, Lg. 13.77-9.
 Review of *Thesaurus linguae latinae epigraphicae* 2.2, Lg. 13.79; 2.3-4, Lg. 13.253-4.
 Review of J. B. Hofmann, *Lateinische Umgangssprache*², CW 30.205-6.
 Review of T. Hudson-Williams, *A short grammar of Old Persian*, JAOS 57.193-4.
 Review of W. Brandenstein, *Die erste 'indogermanische' Wanderung*, AJP 58.240-2.
 Review of A. Walde, *Lateinisches etymologisches Wörterbuch*³, revised by J. B. Hofmann, fasc. 9, AJP 58.372-3.
 Review of H. Hirt, *Indogermanische Grammatik VII: Syntax II*, Lg. 13.246-9.
 Review of E. Benveniste, *Origines de la formation des noms en indo-européen*, Lg. 13.249-52.
 Review of J. Svennung, *Kleine Beiträge zur lateinischen Lautlehre*, CW 31.30-1.
 Review of H. Koppelman, *Die eurasische Sprachfamilie, Indogermanisch, Koreanisch, und Verwandtes*, Lg. 13.321-3.
 Review of H. Frisk, *'Wahrheit' und 'Lüge' in den indogermanischen Sprachen*, Lg. 13.323-4.
 Review of L. Brenner, *Entwicklung der Funktionen der lateinischen Konjunktion dum*, Lg. 13.325-6.
 Review of H. Bäck, *The synonyms for 'child', 'boy', 'girl' in Old English*, Lg. 13.325-6.
 Review of M. Pallottino, *Elementi di lingua etrusca*, Lg. 13.331-2.
 Review of F. Butavand, *Le secret du texte étrusque de la momie de Zagreb*, Lg. 13.333.
 Review of J. Duchesne-Guillemin, *Études de morphologie iranienne I: Les composés de l'Avesta*, JAOS 57.329-31.
1938 Varro on the Latin language, with an English translation (Loeb classical library), two vols., pp. 1, 676; Cambridge and London.
 The restoration of order by Darius, JAOS 58.112-21.
 Old Persian jottings, JAOS 58.324-30.
 Review of D. Jones, *An English pronouncing dictionary*⁴, Lg. 14.78-9.
 Review of E. Weekley, *Surnames*³, Lg. 14.79-80.
 Review of W. Fenzlau, *Die deutschen Formen der litauischen Orts- und Personennamen des Memelgebiets*, Lg. 14.80-1.

- Review of J. Laukis, *Lithuanian self-instruction*, and Kišeninis, *Zodynėlis Lietuviškai-Angliškai ir Angliškai-Lietuviškai*³, Lg. 14.81-2.
- Review of M. A. Simsar, *Oriental manuscripts of the John Frederick Lewis Collection in the Free Library of Philadelphia*, Lg. 14.160.
- Review of G. Devoto, *Tabulae iguvinæ*, Lg. 14.212-8.
- Review of P. S. Costas, *An outline of the history of the Greek language, with particular emphasis on the koine and subsequent periods*, CP 33.319-20.
- Review of S. Chase, *The tyranny of words*, CW 32.15-7.
- Review of J. Whatmough, *The foundations of Roman Italy*, CP 33.227.
- Review of L. F. Sas, *The noun declension system in Merovingian Latin*, CW 32.65-6.
- Review of A. O. Grubb, *French sports neologisms*, Lg. 14.300.
- Review of H. H. Carter, *Palaeographical edition and study of the language of a portion of Codex Alcobacensis*, Lg. 14.300-1.
- Review of J. Mersand, *Chaucer's Romance vocabulary*, Lg. 14.301-2.
- Review of H. Lewis and H. Pedersen, *A concise comparative Celtic grammar*, Lg. 14.304-5.
- Review of F. W. König, *Relief und Inschrift des Königs Dareios I am Felsen von Bagistan*, JAOS 58.675-9.
- 1939 The Nakš-i Rustam inscriptions of Darius, Lg. 15.160-77.
Reconstructing the history of a language, CW 32.244-8.
Review of E. Herzfeld, *Altpersische Inschriften*, JAOS 59.126-9.
Review of I. Goldberg, *The wonder of words*, Lg. 15.121-3.
Review of A. C. Juret, *Formation des noms et des verbes en latin et en grec*, Lg. 15.125-7.
Review of A. C. Juret, *La phonétique latine*²; *Principes de métrique grecque et latine*²; *Phonétique grecque*, Lg. 15.127-9.
Review of V. Väänänen, *Le latin vulgaire des inscriptions pompéiennes*, AJP 60.382-4.
- 1940 *The sounds of Latin*, 2d ed., revised (Special publication of the LSA), pp. 220; Baltimore: LSA.
Review of A. Ernout and A. Meillet, *Dictionnaire étymologique de la langue latine*², Lg. 16.57-9.
Review of A. Perrot, *Malédiction et violations de tombes*, AJA 44.153.
Review of A. Walde, *Lateinisches etymologisches Wörterbuch*³, revised by J. B. Hofmann, fasc. 11, AJP 61.513-4.
- 1941 The Greek aspirated perfect, Lg. 17.189-93.
Review of E. Schwyzler, *Die Parenthese im engern und im weitem Sinne*, Lg. 17.71-3.
Review of H. Adank, *Essai sur les fondements psychologiques et linguistiques de la métaphore affective*, Lg. 17.73-4.
Review of L. H. Gray, *Foundations of language*, CW 34.249-51.
Review of E. H. Warmington, *Remains of Old Latin III*, CJ 36.559-63.
Review of M. Lejeune, *Observations sur la langue des actes d'affranchissement déphiques*, Lg. 17.164-6.
- 1942 Vocalic *r* in Old Persian before *n*, Lg. 18.79-82.
Old Persian Texts, I. The Darius Suez c inscriptions; II. An inscription of Darius II, JNES 1.415-23.
Old Persian Studies, JAOS 62.266-77.
Review of S. Sen, *Old Persian inscriptions of the Achaemenian emperors*, Lg. 18.154-6.
Review of *Collectanea Schrijnen*, AJP 63.487-8.
- 1943 Addendum on Bh. 1.86; Addendum on Bh. 4.44, JAOS 63.67.
Old Persian Texts, III. Darius' Behistan inscription, Column V, JNES 2.105-14.
Studies in Old Persian morphology, Lg. 19.221-9.
The etymology of Latin *sine*, TAPA 73.54-7.
Old Persian Texts, IV. The lists of provinces, JNES 2.302-6.
Review of J. Brunel, *L'aspect verbal et l'emploi des préverbes en grec, particulièrement en attique*, AJP 64.121-2.

- Review of E. H. Sturtevant, *The Indo-Hittite laryngeals*, Lg. 19.165-8.
 Review of E. H. Sturtevant, *The pronunciation of Greek and Latin*², AJP 64.379-81.
- 1944 The Old Persian relative and article, Lg. 20.1-10.
 Old Persian Texts, V. Darius' Behistan inscription, Column V: A correction, JNES 3.332-3.
- 1945 *The sounds of Latin*, 3d ed., revised (Special publication of the LSA), pp. 220; Baltimore: LSA.
 Old Persian Texts, VI, Darius' Naqš-i-Rustam B inscription, JNES 4.39-52.
 The name of Hystaspes, Lg. 21.55-8.
 Old Persian Texts, VII. Artaxerxes I, Persepolis A; VIII. Addenda On Naqš-i-Rustam B; X. Naqš-i-Rustam D, JNES 4.228-33.
 Old Persian *artācā brazmaniya*, Lg. 21.223-9.
- 1946 *The forms of Latin: A descriptive and historical morphology* (Special publication of the LSA), pp. 159; Baltimore: LSA.
 The accusative in Old Persian *mām kāma*, JAOS 66.44-9.
 The oldest Old Persian inscriptions, JAOS 66.206-12.
 Review of F. Bodmer, *The loom of language*, CJ 41.180-1.
 Review of P. P. Bagan, *The syntax of the letters of Pope Gelasius I*, CP 41.183-4.
 Review of J. Humbert, *Syntaxe grecque*, CP 41.234-7.
 Review of *Mélanges de philologie, de littérature et d'histoire anciennes offerts à Alfred Ernout*, Lg. 22.261-3.
 Review of D. Norbert, *Syntaktische Forschungen auf dem Gebiete des Spätlateins und des frühen Mittellateins*, CP 41.237-8.
- 1947 On some Old Persian inscriptions of Darius I, JAOS 67.30-33; correction 67.152.
 On the pronunciation of vowel quantities in Latin, CW 41.91.
 Review of P. Chantraine, *Morphologie historique de grec*, Lg. 23.56-8.
 Review of P. Chantraine, *Grammaire homérique (phonétique et morphologie)*, AJP 68.107-8.
 Review of *Mémorial des études latines ... offert ... à J. Marouzeau*, AJP 68.108-9.
- 1948 Addenda on Varro, *De lingua latina*, TAPA 68.123-30.
 A problem of Latin prosody, *Mélanges Marouzeau* 303-8.
 An unpublished letter of Oliver Wendell Holmes, *American literature* 20.333-6.
 Pedagogue's progress, CW 42.82-7.
 Review of E. H. Sturtevant, *An introduction to linguistic science*, CW 41.216-9.
 Review of V. Pisani, *Linguistica generale e indoeuropea: Saggi e discorsi*, Lg. 24.194-5.
 Review of M. Lejeune, *Traité de phonétique grecque*, Lg. 24.195-8.
 Review of W. Hinz, *Altpersischer Wortschatz*, JAOS 68.151-3.
 Review of *Eranos Rudbergianus*, CW 42.63-4.
- 1949 Review of A. Walde, *Lateinisches etymologisches Wörterbuch*², revised by J. B. Hofmann, Vol. 2, pp. 1-160, Lg. 25.53-6.
 Review of V. Pisani, *Introduzione alla linguistica indoeuropea*, Lg. 25.196-8.
 Review of V. Pisani, *Grammatica latina storica e comparativa*, Lg. 25.198-203.
- 1950 *Old Persian grammar, texts, lexicon* (American oriental series, Vol. 33), pp. xiv, 216; New Haven: American Oriental Society.
 Review of K. Gries, *Constancy in Livy's Latinity*, CW 43.203-4.
 Review of *Mélanges de philologie, de littérature et d'histoire anciennes offerts à J. Marouzeau par ses collègues et élèves étrangers*, Lg. 25.204-6.
 Review of A. Walde, *Lateinisches etymologisches Wörterbuch*², revised by J. B. Hofmann, Vol. 2, pp. 161-352, Lg. 25.306-10.
- 1951 Cameron's Old Persian readings at Bisitun: Restoration and notes, *Journal of cuneiform studies* 5.55-7.
 Review of H. Huisintveld, *De populaire elementen in de taal van M. Valerius Martialis*, CW 44.107.
 Review of J. B. Hofmann, *Etymologisches Wörterbuch des Griechischen*, 1. Teil, AJP 72.79-81.

INDO-EUROPEAN NEGATIVE COMPOSITION

JAAN PUHVEL

Harvard University

1. FUNCTIONAL THEORY OF INDO-EUROPEAN **η*-

The functional history of the IE negative prefix, the variants of which we conveniently designate by their central form **η*-, is an integral part of the history of IE composition. A classification of IE compounds, based on historical rather than logical considerations, would include (1) old types, generally declining during the historical period of IE—bahuvrīhis and synthetic compounds (verbale Rektionskomposita), and (2) distinctly younger types—pronominal hypostatics, determinatives, and copulatives.

The earliest types available to us in which **η*- appears are bahuvrīhis and (synthetic) verbal adjectives.¹ A thorough analysis of these types is therefore essential for an adequate comprehension of the nature of **η*-.

The bahuvrīhi type has been variously explained. Whereas in earlier times a mechanical consideration of its semantic contents led to the assumption that it represents a secondary development from the determinatives with added possessive value and conversion to adjectives (e.g. Brugmann, KZ 24.40), more modern theories² have made unquestionably clear the primitive and primary nature of the type. In the most primitive forms of IE—Vedic, Gathic, Homeric Greek, Gothic—the bahuvrīhis are found in their greatest expansion, whereas the tatpuruṣa type is rare and still largely expressed by uncompounded elements; its gradual extension can be followed in the later stages of the languages.³

The synthetic compounds⁴ form, together with the bahuvrīhis, the oldest layer, and seem to have been the most common primitive type. A survey by Risch (IF 59.5) of the Greek material, which has a recognized conservative character, shows that about 60% of all compounds come under the heading of synthetics, 25–30% are bahuvrīhis, and the rest are mainly hypostatics, determinatives, and occasional copulatives.

The synthetic type comprises variants with the verbal element second (Gk. *ἄζυξ*, Skt. *ayuj*, Lat. *coniux*) and with the verbal element first (Gk. *ἐλκεσίτων*, *ἐλκεσίπεπλος*, Skt. *vidād-vāsu*).⁵ This classification has only lately been recognized. The whole *ἄζυξ* type was classed with the tatpuruṣas by the Sanskrit grammarians, which is logically acceptable but historically impossible.

¹ Hamilton, *The negative compounds in Greek* 27 (Baltimore, 1899); Frisk, *Gebrauch des Privativpräfixes im indogermanischen Adjektiv* 4, 45 (Göteborg, 1941).

² Jacobi, *Compositum und Nebensatz* 83 ff. (Bonn, 1897); Brugmann, IF 18.59 ff.; Petersen, IF 34.249 ff.

³ Hirt, *Indogermanische Grammatik* 4.67 ff. (Heidelberg, 1928); Frisk, IF 52.288; Petersen, IF 34.273; Risch, IF 59.1 ff., 245 ff., 292 f.; Schwyzler, *Griechische Grammatik* 1.453 (Munich 1939).

⁴ Jacobi 1 ff.; Hirt 4.25 ff.

⁵ The second type concerns us only indirectly, since a priori it contains no compounds with the negative prefix.

The chief characteristic of the synthetic type is the concrete character of the verbal element and its frequent non-existence or secondary appearance outside the compound. The oldest type shows the bare verbal root ($\acute{\alpha}\zeta\nu\acute{\xi}$), which was, however, extended already in IE times: by $-t$ in Gk. $\acute{\alpha}\delta\mu\acute{\eta}\varsigma$, Skt. $i\check{s}ubhrt$, Lat. *comes*; by $-o-$ in Gk. $i\pi\acute{o}\delta\alpha\mu\omicron\varsigma$, Skt. *abhram-liha*, Lat. *causidicus*. Besides, the separate languages exhibit numerous individual increments. Common to Greek and Latin is the type in $-ā$: Gk. $\delta\epsilon\sigma\acute{\rho}\omicron\tau\eta\varsigma$, $\pi\omicron\lambda\iota\pi\acute{o}\rho\theta\eta\varsigma$, Lat. *indigena*, *aestuma*.

In these examples the concrete verbal elements appear exclusively in composition. Uncompounded parallels, like most radical nouns (e.g. Gk. $\sigma\acute{\tau}\acute{\upsilon}\xi$, $\delta\rho\acute{\alpha}\xi$, Lat. *nex*) show an abstract character in opposition to the compounds:⁶ Gk. $\phi\acute{\upsilon}\gamma\alpha\delta\epsilon$ 'to flight', $\pi\rho\acute{o}\sigma\phi\upsilon\acute{\xi}$ 'fugitive'. This differentiation in use is partly broken through; in Sanskrit, although concretes like $-lih$ and $-bhuj$ are restricted to compounds, many verbal roots can also be used as uncompounded concretes: *dviṣ* (masc.) 'enemy', (fem.) 'hostility'. Similarly Gk. $\pi\tau\acute{\omega}\xi$, $\phi\acute{\omega}\rho$, whereas Lat. *rēx*, *dux* are borderline cases of this secondary development. Lat. *rēx* and OIr. *rī* are found for the most part independently, but in other languages the radical form shows a tendency to appear in composition (Gaul. *Dumnorix*, Skt. *samrāj*), whereas the simplex shows extended forms (Skt. *rājā*). The same is seen in Lat. *dux*, German *Herzog*.

The appearance of the bare root has considerable historical significance. In the complex morphological structure of pre-migrational IE, a simple root was out of place. This is also borne out well by the general decrease of radical nouns in the various IE dialects; Balto-Slavic and Germanic have eliminated the type almost completely. It seems that we have to do with relics from an early period before the rise of the IE inflexional system. The original value of the bare root must be analyzed accordingly. To assume that it functioned as a complicated kind of relative participle⁷ (i.e. that the root alone was capable of assuming complicated grammatical functions) is to deprive inflexion of any *raison d'être*. We assume instead that the root had no more definite value than its appearance indicates: it is a relic from a time when uninflected isolating juxtaposition reigned in IE, not unlike certain tendencies in modern analytical languages like English. Thus the root expressed the idea pure and simple, and only its use would indicate its grammatical relations. If used alone as a subject, it expressed the basic verbal idea, being thus a primitive action noun. It could be qualified by another noun in juxtaposition, and the whole expression, similarly 'given as name', could be applied appositionally to a noun as a *bahuvrihi*.

If this analysis is correct, it sheds light on the essence of early IE composition: the primitive compounds seem to point to an original isolating type of language; when more synthetic grammatical structure developed, remains of the former type assumed the function of compounds through the fossilization of the juxtaposed elements and not by any active process of composition. Therefore they are radically different from the structure of an inflexional language, and generally decline in later IE. For the same reason they cannot be divided into their

⁶ Chantraine, *La formation des noms en grec ancien* 4 (Paris, 1933).

⁷ Jacobi's explanation of synthetic composition.

elements by a mere process of mechanical division, since they were not compounded on the later principles of determinatives and the like.

Gradually, through the general decrease of un-compounded radical abstracts, the bahuvrīhi conception was obscured, and through direct association with the verb the synthetic forms came to be felt as a kind of compounded agent nouns or verbal adjectives. Hence direct creation in this sense was possible, and the synthetic elements could also appear in similar un-compounded use. There are thus two reasons why many of the synthetic verbal elements exist only in compounds: the disappearance or replacement of the primitive abstracts, and the expansion of the compounded bahuvrīhi concretes.

This conception would show the original close connection between the synthetic compounds and the bahuvrīhis by demonstrating the necessity of the original bahuvrīhi conception of both types and thereby the uniform origin of early IE composition and the use of **η*-.

What has been said about the root type applies also to the extensions -*t*-, -*o*-, and -*ā*; it seems likely that abstract (or collective) value is original in all of them.

The abstract is seen in Skt. *stut* 'praise, hymn', *samit* 'meeting', but this type is generally extended to -*tis* (ἀγνώς : γνῶσις, *iṣubhrt* : *bhrtis*). That a concrete sense developed secondarily, through bahuvrīhi intermediary, is made probable by frequent exclusive appearance in compounds and by the very vagueness of the meaning, which can show rather indifferently active or passive value, exactly as in the bare root type (νῆ(ς)ις 'unknowing', ἀζυγέ 'unyoked, unmarried'). In Sanskrit, the active value predominates (*nāmadhās*, *iṣubhrt*), and the passive value is rare (*devaṣrut*); in Greek, on the other hand, the passive value is prominent, but many forms have a double sense: ἀγνώς 'unknown' and 'unknowing'; ἡμιβρώς 'half-eaten', but ὠμοβρώς 'eating raw meat'; ἀκμής 'untiring', but δουρικμής 'slain by a spear'.⁸

The same original vagueness of the concrete meaning is seen in the thematic type. Gk. *τόμος* means 'a cut'; in a concrete sense it appears in composition in bahuvrīhi use, e.g. *λιθοτόμος* 'associated with stone-cut', a meaning which can develop into active or passive. In this common type accentual differentiation is essential to avoid serious handicaps of comprehension: *λιθοτόμος* 'cutting stone' : *ἀτομος* 'not cut'. From such opposition resulted the non-synthetic use as agent nouns: Gk. *τόμος* 'cut' : *τομός* 'cutter', Skt. *vāras* 'choice' : *vārās* 'suitor'.

The -*ā* type is originally abstract-collective: Gk. *τομή*, *φύγη*, Lat. *fuga*. That agent-noun value is secondary here is indicated by its absence from Sanskrit and by its compounded character, which points to original bahuvrīhi conception (Lat. *trānsfuga*, *aestuma*).

The verbal adjectives in -*tos* are not generally considered synthetic, but a number of striking correspondences with the other types may warrant a similar interpretation. This is particularly important since these verbal adjectives are among the earliest and most frequent compounds with **η*-.⁹

IE -*to(s)* is the thematic variant of the root extended by -*t*-, and the two

⁸ Buck, Cl.Ph. 12.173 ff.

⁹ Frisk 4 f.

types are found side by side (Gk. *προβλῆς* : *πρόβλητος*, *ἀγνώς* : *ἄγνωτος*). Uncompounded forms (Gk. *θάνατος*, Skt. *ghātas*, Lat. *lectus*) are used as abstracts. This is old, as is shown by ablaut differences. It seems that the accent was on the root syllable in the use as abstracts, but on the ultima in corresponding verbal adjectives. Thus to IE **k̑leutom* 'a hearing', Avestan *sraotem*, Gothic *hliuþ*, corresponds the Skt. concrete *ḡrutás* 'heard', and Greek has the alternation *θάνατος* : *θνητός*. Parallel to *-ā* there are feminine abstracts in *-tā* (Gk. *πινυτή*, Lat. *multa*),¹⁰ and the general agreement of *θάνατος*, *θνητός*, *πινυτή* with *τόμος*, *τομός*, *τομή* is complete. Another fact illustrates the connection between the *-tos/-tā* types and those previously analyzed: where other dialects like Indo-Iranian and Italic concretized abstracts in *-t*, Greek shows parallel forms with *-t-ā*: *-φόντης*, Skt. *-hat*; *ὤμησθης*, Skt. *āmād*; *ἐπιστάτης*, Lat. *antistes*.

This general parallelism of *-tos* with the other types leads us to the assumption of similar bahuvrīhi origin for the verbal adjectives in *-tos*.

Languages where the *-tos* formation has been drawn into the verbal system with subsequent generalization and expansion, as Sanskrit and Latin, are of little value here, whereas the conservative character of Greek and particularly Homeric usage have to be given most consideration. In Homer the verbal adjectives in *-tos* are found to occur mainly in composition;¹¹ many of the forms are barred from uncompounded use, e.g. *-κμητος*, *-δάϊκτος*, *ἄϊστος*, *ἄελπτος*, *ἄφθιτος*. The Rig-Veda has only *ākṣitas*, Latin only *invītus*. Vagueness of the original meaning is seen here as well. The predominant sense has become a past and passive one, but this is mainly due to the widespread incorporation of the type into the verbal system. Active value is seen in *ὀδυνήφατος* 'pain-killing' (contrast *Ἀρηϊφατος* 'slain by Ares'), *ἄϊστος* 'unknown' or 'not knowing', *ἀνόητος* 'not understanding'. Important is the fact that Greek compounds with *-τος* are 'two-ending adjectives' like the bahuvrīhis and the thematic synthetics, i.e. the last element still shows the generic inflexibility of a noun. This type is distinctly older than the general uncompounded three-ending variety.

Greek has the twofold type *ἀκάματος*, *ἄκμητος*. Although historical development differentiated their makeup, both reflect a bahuvrīhi; but *ἀκάματος* retains a closer bahuvrīhi feeling and more direct association with *κάματος* because of phonetic similarity. This conception is supported by the parallelism of *ἀκάμας* : *ἀκμής*.

There are in Greek genuine bahuvrīhis with the *-tos* suffix, the latter adding nothing to the meaning: *ἀκήδεστος* = *ἀκηδής*; *ἀτέλεστος* = *ἀτελής*. In Sanskrit we find *anaptas* 'without water', in Latin *insegestus* 'not sown', *intempestus* 'unseasonable'. This seems to point to a time when the bahuvrīhi notion was still vividly felt in the verbal adjectives and was thereby conducive to the transfer of *-tos* to purely nominal bahuvrīhis as a characteristic suffix like Skt. *-ka* (*amanas-ka* 'ἄφρων').

IE *-tos* is used to form adjectives from nouns, giving the sense of 'provided with': Gk. *πρεπώτος*, *θαυματός*, Lat. *honestus*, *barbātus*. Lat. *laudātus* and *partītus* belong strictly to the same type, which could thus be described as a 'denominative

¹⁰ Carruthers, Cl.Ph. 26.181 ff.

¹¹ Buck-Petersen, *A reverse index of Greek nouns and adjectives* 469 (Chicago, n.d.).

participle'. But here also the connection with the verbal system is purely secondary. The type has its independent origin, which is perhaps best sought in an abstraction from a compounded bahuvrīhi type, where the suffix naturally came to possess the meaning 'provided with': ἀγένητος = ἀγενής.

On IE abstracts are based two remarkable types of Greek synthetics, both of which are individual Greek innovations. It is noteworthy that we here see bahuvrīhi character; it seems highly likely that we are witnessing processes closely related to the one we have assumed for IE. The first is the type in -ής (ἀφανής, ἀσεβής). The suffix -ής is the common IE ending of bahuvrīhis which contain s-stem neuters as their second element (Gk. μένος : εὐμενής; Skt. *manas* : *śumanās*).

It is probable that the following were originally bahuvrīhis of the same kind: ἀσεβής (σέβας), ἀδεής (δέος), ἀβλαβής (βλάβος). By uncompounded parallels of more immediate verbal nature, which in this case happen to be their logical opposites—σεμνός, δειλός, βλαβερός—they were drawn into a verbal semantic circle; as a result, new forms could be created directly from verbal roots after the pattern of the existing ones, e.g. ἀφανής, ἀσφαλής in opposition to φανερός, σφαλερός; ἀσυνής, ἀτειρής, ἀτυχής.

The other suffix, -μων, corresponds closely to -ής, being the long-grade bahuvrīhi ending of n-stem neuter abstracts: πρᾶγμα : ἀπράγμων. The type was conceived as synthetic verbal adjectival, and new instances were created directly: ἀφράδμων, ἀδαήμων, ἀφραδής, ἀδαής without any *φράδμα, *δάημα or *φράδος, *δάος. The isolated νήστις may be a synthetic based on an old -tis abstract.

We have attempted to trace the consistent development of the synthetic type of composition out of primitive action nouns and its original bahuvrīhi character, hoping thereby to demonstrate the uniform nature of early IE composition and also, particularly, of the original use of the negative prefix. We thus assume that *n- in primary use was originally confined to bahuvrīhis. The attested varieties are as follows:

1. Ordinary bahuvrīhis:

- 1a. without change of ending: Gk. ἄθεος; Skt. *adevas*
- 1b. with thematizing (adjectivizing) increment in consonant stems:
ἄνδρος; *anudras*
- 1c. with generic adaptation of ending: ἄτιμος (τιμή); *aṣaṇkas* (ῥαῆκᾱ)
- 1d. with ablaut modification of ending: ἀτελής, ἄφρων; *atejās*
- 1e. with change of increment (other than generic; possibly original adjectivization): ἀναλκις (ἀλκή); Lat. *imbellis* (*bellum*)
- 1f. with addition of increment (other than thematizing): ἀκήδεστος;
anaptas, *amanaskas*; Lat. *insegestus*
- 1g. with change or addition of increment and thematization: ἀτέλε(σ)ιος;
aputriyas

2. Synthetic bahuvrīhis:

- 2a. bare root: ἄζυξ; *ayuj*
- 2b. -t increment: ἀδμής
- 2c. thematized root: ἄτομος
- 2d. thematized -t increment: ἄφθιτος; *akṣitas*; Lat. *invītus*

Parallel creations were present in the separate dialects (e.g. Gk. *ἀδαής, ἀδαήμων*). A spread from synthetic bahuvrīhis to primary participles or verbal adjectives may have occurred already in IE (Skt. *divdvān*, Lat. *īnsciēns*, Goth. *unwītands*). Later developments were manifold and varied. The Indic preference for composition and negative expression brought about a wide expansion of the use of **η-*, already seen in the Rig-Veda, a spread to verbal adjectives and participles in general and to practically every kind of nominal expression. Homeric Greek shows that remarkable restraint which is our main source for the investigation of IE conditions, and even in later Greek the expansion is slow and gradual.¹²

2. NATURE OF THE NEGATIVE PREFIX

We distinguish the following 'negative categories':

1. negative proper: the sentence negative IE **ne*
2. privative: **η-* in bahuvrīhis
3. reversive: **η-* and **ne* in other categories

Between the use of **η-* in the primitive compositional types (ordinary and synthetic bahuvrīhis) and its use in determinative composition (karmadhāraya) there is the same radical difference in kind as between the types themselves. For conclusions about the primitive nature of **η-* we have to consider only the primitive types; the common later use in determinatives is secondary, developed through transfer of **η-* from the older types, much as the adverbial element **dus-* is carried over from the bahuvrīhis (*δυσδαίμων, durmanās*) to newer composition (*Δυσπαρις, durjanas*).

The value of **η-* in bahuvrīhis is privative or separative. Thus, *ἄπαις* means 'no-child-having, without child', or rather 'child not there'; the latter would seem rather close to the purely negative meaning that we ascribe to the sentence negative **ne*. However, there is a radical functional difference: **ne* is foreign to bahuvrīhis, even in Balto-Slavic where it is generally used in composition. These facts refute the contention that there is no very deep difference between sentence and word negatives,¹³ and even induce doubts about any original connection between **η-* and **ne*.

In the newer types, on the other hand, the negatives form a compositional element which contributes to bring about a compounded notion through synthesis with another element. This new notion is in its essence positive. The whole expression does not basically denote non-existence or non-activity; rather, it states that the original semantic content of the second element is excluded.

The need for the definition of such new notions in compounds is present also in verbs. Whenever **ne* comes to form a close semantic link with a verb and unites with it into a single notion, the verbal compound is of the same nature as the determinative compounds with **η-*: we have a notion defined through exclusion, not a denial of a notion. As *unholy* is about the same as *profane*, so Gk. *οὐκ ἐώ* is 'I forbid', and *οὐ φημι* is 'I deny'; similarly OSl. *velēti* 'bid', *nevelēti* 'forbid'.¹⁴

¹² Cf. Frisk for Greek, Indo-Iranian, Latin, Germanic; Hamilton for Greek.

¹³ E.g. Hirt 4.49, 5.390, 7.69.

¹⁴ Cf. Miklosich, *Vgl. Grammatik der slavischen Sprachen* 4.173 (Vienna, 1868-74); Delbrück, *Vgl. Syntax der indogermanischen Sprachen* 2.522 f. (Strassburg, 1897).

Both these uses are secondary and closely parallel. In both, the negative has undergone a transfer and a kind of semantic mutation. From bahuvrīhi, **η*- is expanded to determinatives through the intermediary of the synthetic variety, and acquires a value which tends more and more to obscure the original sense because of the gradual decay of the bahuvrīhi type; **ne* is transplanted to verbs as a word negative, and undergoes a semantic shift by being transferred from its stressed sentence position into the equilibrium of semantic stress which characterizes a compound.

In nominal composition, **ne* is found sporadically in the newer use (determinative composition or synthetics with faded bahuvrīhi character, e.g. Skt. *na-puṃsaka*, Lat. *nefās*, *nescius*), which seem to result from semantic combination similar to the one seen in verbs, or from late analogical substitution.

We shall now attempt to analyze the intrinsic nature of **η*- in its primitive bahuvrīhi use.

Beside the general category of bahuvrīhi with nominal first element, authoritatively explained by Petersen's general principle of primitive name-giving, there is a slightly different group, the *ἐνθεος* type.¹⁵ The general explanation for this is the one that Jacobi sought for bahuvrīhi in general, namely 'Univerbierung' from a time when the verb 'to be' was merely implied in appositional expressions. This explanation is accepted for the *ἐνθεος* type even by the opponents of Jacobi's general approach (e.g. Petersen, IF 34.282). The type is thus explained as originating from primitive elliptic clauses in apposition, which is actually the same as the original isolating juxtaposition or 'name-giving' that we have assumed previously: Gk. *φρήν ἐνθεος* 'a mind, therein God'; *κρίως ἐπερος* 'a ram, thereon wool'; Skt. *naro vyaṅgaḥ* 'a man, limb(s) off'.

It seems likely that the **η*- type belongs to the same category; *ἄθεος* seems to be the opposite of *ἐνθεος*. Taking **η*- as an adverbial element of original separative value is the most natural way to account for the **η*- bahuvrīhi. Here Balto-Slavic offers useful evidence. This group has eliminated **η*- (except for relics like OS. *q-rodŭ* 'ἀφρων'), and uses in composition the sentence negative **ne*. The process can be reconstructed thus: the elimination of **η*- dates from a time when the bahuvrīhi use had already been overshadowed by determinatives; Balto-Slavic went one step further and began using reversive **ne*- in nominal expressions as in verbs. Bahuvrīhi with *ne*- are practically non-existent; some may have assumed it (e.g. OS. *ne-bogŭ* and perhaps *neplody*) or retain traces of original **η*- (*qrodŭ*). But mainly we find instead parallels of the *ἐνθεος* type with more vivid concrete adverbs of separative value: OS. *bezbogŭ* 'ἄθεος', *bezdrokŭ* 'ἄχειρ', *ubogŭ* 'poor', cf. *nebogŭ*. Vondrak¹⁶ claims these forms to be hypostatic (*bez boga* > *bezbogŭ* as *ἐν κεφαλῇ* to *ἐγκέφαλος*); but this is unlikely if we consider the probable antiquity of the type and the original value of *bez* as a separative adverb (Skt. *bahis* 'out, forth, off'). It is, rather, probable that the hypostatic conception was secondarily felt to be inherent in the type when *bez* had come into use as a preposition 'without'; OS. *člověkŭ bezbogŭ* 'man, God off' > 'man without God'; similarly Gk. *ἀνὴρ ἄθεος* 'man, God off' > 'man, no-

¹⁵ Brugmann, IF 18.127.

¹⁶ *Vgl. slavische Grammatik* 1.676 (Göttingen, 1906-08).

God(-having)'. In *u-bogŭ*, *u-* has been compared with Skt. *ava*, Gk. *οὐ*¹⁷ as containing a similar separative adverb. Güntert (IF 40.186 ff.) makes an ingenious and complicated attempt to demonstrate the origin of Gk. *οὐ* from the adverb Skt. *ud* 'up, forth' and to prove the semantic aspects by similar examples. We cannot accept Güntert's approach, since it fails to demonstrate how *οὐ* could become the predominant sentence negative from bahuvrīhis of the *ἐνθεος* type and disappear again from bahuvrīhis, and why it did not instead supersede the negative prefix. Still, Güntert has indicated an approach which might prove more fruitful when applied to IE **h*-.

The use of separative adverbs in bahuvrīhis is common throughout IE where the bahuvrīhi is still alive. Besides the Balto-Slavic cases the following types illustrate the point: Skt. *nirdevas* 'God off', *adevas*; *vihasas* 'hand(s) off', *ἄχειρ*; Lat. *expers* 'ἀμερής', *excors* 'ἄφρων', *exsomnia* 'ἀπνους', *exlex* 'ἀνομος'.¹⁸ Gk. ἀποφώλιος 'irritus' may be a similar bahuvrīhi. Germanic examples are Goth. *uswēna*, OE *orwēne*, OHG *urwāni* 'exspes', OE *orblēde*, OHG *urploti* 'exsanguis', OHG *urherzi* 'excors'. There are besides, in OE and OHG, a group of words which show a negative prefix *æ-* and *ā-* respectively.¹⁹ These point to PGmc. *æ* and IE *ē*; they are probably instances of the adverbial root **ē*.²⁰ This explanation is made particularly probable by the bahuvrīhi character of several of the words: OE *æ-menn* 'ἀνανδρος', *æ-not* 'ἀχρεῖος', OHG *ā-teilo* 'ἀμερής', *ā-wiggi* 'ἀνοδος', Lango-bardian *ā-mund* 'without guardian, free'. As bahuvrīhis are found in relics only in the old Germanic dialects (except Gothic), these cases clearly illustrate the primitive character of the type.

We assume that a similar practice was present in IE where an adverbial root of separative meaning was used in bahuvrīhis. It is thus suggested that the affinities of the negative prefix are to be sought among IE concrete adverbial roots, perhaps the widespread group illustrated by Avest., Goth., OHG *ana*, Gk. *ἀνα*, Slav. *na*, with the original meaning 'along, forth, off'.²¹

3. FORM OF THE NEGATIVE PREFIX

The form of the negative prefix is intimately connected with two important features of early IE phonology: the loss of laryngeals and the allophonic variation of resonants.²²

¹⁷ Meillet-Vendryes, *Traité de grammaire comparée des langues classiques* 545 (Paris 1927).

¹⁸ A later development seems to have confused the original picture, especially in Greek. The *ἐνθεος* bahuvrīhis date from a time when *ἐν* etc. were purely adverbial expressions, as still seen in Homer: *ἐκ δ' ἄγαγε κλισίης Βρισηίδα*, where the ablative genitive is self-sufficient and *ἐκ* merely strengthens it. With the development of prepositional value, many of the bahuvrīhis could be interpreted as hypostatics, like Slavic *bez*. It is uncertain in cases like *ἐκτιμος* and *ἐκπνους* 'out of breath' whether we have to do with original bahuvrīhis or with later hypostatics.

¹⁹ Fowler, *The negatives of the Indo-European languages* 32 (Chicago, 1896); Gray, *Lg.* 1.127.

²⁰ Walde-Pokorny, *Vgl. Wb. der indogermanischen Sprachen* 1.95 (Berlin und Leipzig, 1930); Pokorny, *Indogermanisches etymologisches Wb.* 280 (Bern, 1949-).

²¹ Walde-Pokorny 1.58; Pokorny 39.

²² Cf. in general Sturtevant, *Indo-Hittite laryngeals* (Baltimore, 1942); Lehmann, *Proto-Indo-European phonology* (Austin, 1952).

In our discussion, the following symbols are used:

<i>S</i>	syllabic
<i>N</i>	non-syllabic
<i>O</i>	obstruent (incl. <i>s</i>)
<i>H</i>	laryngeal
/	pause

The allophones of the resonants (semivowels) have been described by Edgerton (Lg. 19.83 ff.).

If the negative prefix is connected with *ἀνα* etc., we must assume an original *a*-colored laryngeal (Hitt. *ḫ*(*h*) or nil). In any case consistency requires that some laryngeal be posited before an otherwise initial resonant no less than before an otherwise initial vowel; for a resonant is always potentially syllabic and hence vowel-like in character. This consideration dictates the assumption of an initial voiced laryngeal (lost in Hittite), e.g. in Hitt. *ú-it-ta-an* : Gk. *ἔτος*, besides a voiceless laryngeal in *ú-e-ik-zi* : *ἐκών*, *ú-e-eš-tin* : *ἐννῆμι*. It may be preferable to reconstruct the primitive form simply as a laryngeal plus the resonant *n* (**Hn*). As long as the laryngeals were intact, the prefix always appeared in its syllabic form; for it is a corollary of the foregoing statements that every IE word originally had an initial obstruent or laryngeal. Hence we obtain the formulae

<i>NHηO</i>	<i>/HηO</i>	<i>SHηO</i>
<i>NHηHN</i>	<i>/HηHN</i>	<i>SHηHN</i>
<i>NHηHS</i>	<i>/HηHS</i>	<i>SHηHS</i>

The decisive step was the loss of the laryngeals. In the new situation that now arose, the allophonic rules for resonants had their effect: *NHηO* and */HηO* gave *NηO* and */ηO* respectively; *NHηHN* and */HηHN* lost their laryngeals with compensatory lengthening of the resonant, yielding *N̄ηN* and */̄ηN*; *NHηHS* and */HηHS* also lost their laryngeals, and the resonant thereupon assumed the allophonic variant regular between non-syllabic and syllabic (Sievers' law), *NHηHS* yielding *NnS* after a light syllable, *N̄ηnS* after a heavy syllable or pause; */HηHS* became */nS*. In the remaining variants, *SHηO* gave *SnO* if *S* was a vowel; if it was a resonant, the result was *NηO*. *SHηHN* yielded *SnN*, and *SHηHS* yielded *SnS*.

We thus posit the following forms of the negative prefix for IE after the laryngeals had been lost:

- 1a. *N-η-O*, */η-O* (before original non-laryngeal, from *NHηO*, */HηO*; or from *SHηO* with resonant *S*)
- 1b. *S-n-O* (before original non-laryngeal)
S-n-N (before original laryngeal; in case of resonant *N* restricted to position after short syllabic)
- 2a. *N-ηn-S* (after heavy syllable); */N-ηn-S*
- 2b. *N-n-S* (after light syllable); */n-S*
S-n-S
3. *N-̄η-N*, */̄η-N* (before original laryngeal; in Greek, if *N* is resonant, voiced laryngeal only; before *ῥ* not at all)

Analogical extension of types took place: 1a and 1b were positional variants, exactly like 2a and 2b; but 1 was also in allophonic variation with 2. Hence, after the original pattern of resonant allophones had been broken, the syllabic forms 1a and 2a, being more distinctive, tended to prevail at the expense of the non-syllabic variants, and yielded the most familiar forms of the prefix; **η-* is seen in Skt. *a-*, Gk. *á-* Lat. *in-*, Goth. *un-*, OIr. *ē-* or *in-*; **ηn-* in Skt. *an-*, Gk. *án-* Goth. *un-*, OIr. *an-*. That analogy has been highly operative in this, the most prolific type, is shown also by later facts. Thus in Latin, when phonetic development threatened to differentiate the types, the probable higher frequency of the preconsonantal variety *η-* > *in-* led to its generalization at the expense of *ηn-* > *an-*.

There may be remains of the non-syllabic variety also. We find in Greek traces of the sequence *S-n-N* after short syllabic. Several obscure etymologies can be explained by this assumption as fossilized synthetic compounds with the negative prefix.

Gk. *νόθος* 'illegitimate, born out of wedlock', opp. *γνήσιος*, is labeled 'obscure' by Boisacq, 'unerklärt' by Hofmann. We explain *νόθος* as a synthetic based on the root **X_{medh-}* 'lead, lead home, marry'.^{22a} This grade is seen in Skt. *vadhūs*, Avestan *vadū* 'wife'; Lith. *vedi*, OSl. *vedo* 'lead; marry'; here belongs probably Gk. *ἀθλον*, Hom. *ἄεθλον* (**X_{medh-}*) 'prize in a contest; carried off', but not Attic *ἔδνον* (**Y_{med-}*), Hom. *ἔδνα* (**Y_{med-}*), OE *wituma* 'dowry' with different obstruent and laryngeal. The development of *νόθος* was **SH_ηX_{medh}os* > **S_ηmedh_{os}* > **η_{medh}os* (generalization of this grade) > **ν_{medh}os* > *νόθος*.

Such generalizations of two non-syllabic resonants initially in a word were found in the various IE dialects; certain combinations are abundant in Sanskrit (*न्याह, व्याहgas, व्राजति*), and some specimens are discernible in historical Greek: *μνήμα*; *βροτός*; *βλάξ* (Skt. *mlecchati*, Lat. *flaccus*, OIr. *mlén*) beside *μαλακός*;²³ **φρήγνυμι*; **φλάσιος*. The general Greek tendency is to eliminate such combinations, and the cluster *ν_{f-}* must have lost its *f* in very early times; no traces of 'position' survive in Homeric meter. *νόθος* also lost its grammatical character of compound early, and appears as a 'three-ending adjective'. Thus *νόθος* would be a synthetic compound of the type *ἄτομος*, meaning literally 'not associated with wedlock'. Should there be objections to this rendering and insistence on a primitive 'not wed', the etymology would still stand; the semantic shift would be similar to and probably connected with that of *ἠίδεος* ('widowed' > 'single'). It seems that for the Greeks the word had a negative flavor, since it is found in

^{22a} In these reconstructions, *X* = a voiceless *a*-colored laryngeal (Hitt. *h(h)* or nil), *Y* = a voiceless *e*-colored laryngeal (lost in Hittite), *Z* = a voiced *e*-colored laryngeal (Hitt. *h* or nil).

²³ *μαλακός* had an initial voiceless laryngeal, as is shown by 'position' in Homeric meter (*εὐνῇ ἐνι μαλακῇ* Il. 9.618; *αἰεὶ δὲ μαλακοῖσι* Od. 1.56). Its allophonic history in formulae is as follows: 1. /*H_ηmlS* > /*ηmlS*; 2. possibly also to /*m_llS*; 3. *NH_ηmlS* > *N_ηmlS* (after heavy syllable); 4. to *N_ml_lS* (after light syllable); 5. *SH_ml_lS* remains in **η_mμαλακος*; 6. in other IE languages it generally gives *Š_ml_lS*. No traces survive of 1 and 3. Upon spread to position after short syllabic, 2, 4, and 6 developed the allophone *Š_ml_lS* seen in *βλάξ* etc., which was generalized in the dialects. Similar to these is the allophonic origin of *μνήμα*, *βροτός*, **φρήγνυμι* (*ἀρρηκτος*; *τείχος τε ρήξαν*, Il. 12.198), **φλάσιος* (*στήθεσσι λασίοισι*, Il. 1.189).

enumerations of negative compounds (e.g. νόθον γὰρ καὶ ἀνέγγυον καὶ ἀνίερον φήσομεν αὐτὸν παῖδα, Plato, Polit. 461 b).

This seems to us the most plausible approach. We may note, however, that the only previous suggestion for νόθος, a connection with Skt. *andhas* 'blind', described as 'très douteux' by Boisacq, may also have gained in plausibility through the laryngeal theory. Semantically it is supported by σκότιος, and the general parallelism of *ambhas*, *alpas* : νέφος (**Hnebh-* with voiceless laryngeal), λαπαρός is in its favor. As in opposition to these νόθος never 'makes position' in Homeric meter, we should assume a voiced laryngeal and set up **Znodh-*.

Gk. νωθής 'sluggish', has been connected with ὄθομαι, but can also be explained as containing the IE root **Znedh-* 'strike, push, rush': Ved. *vádhati* 'slay', *vadhar* 'deadly weapon'; Skt. *vadhas* 'murder, slayer'; Avestan *vādāya-* 'push off, repel', *vadar* 'weapon'; Gk. ἔνοσις < *ἐνφοσις 'shaking', (φ)ώθew 'push, rush, impel'. νωθής would thus be a bahuvrīhi like ἀβλαβής from a neuter abstract **Zyōdhos* which gave **nyōdhēs* > **νφωθής* > νωθής, or an analogical synthetic like ἀφανής.

Important in this connection are the Greek types νήκεστος, ἀνήκεστος (beside ἀνάκεστος); νώνυμος, ἀνώνυμος; νηκερδής (beside ἀκερδής). They show almost exclusively bahuvrīhis (ordinary and synthetic); besides, most of them are comparatively rare, unproductive forms which are declining and belong mainly to the archaic poetical language. We must take exception to the long current explanation of these types.²⁴ νη- is explained as containing the sentence negative **ne-* contracted with the initial vowel of the second element (νε + ἀκεστος > νήκεστος), the νη- remaining as a type after the elimination of **νε* and being thus extended (νη-κερδής).

Besides the general objections which may be deduced from the preceding development, there are a number of reasons against this assumption. The contraction is questionable; the attribution of analogical character to old bahuvrīhis like νηκερδής is not convincing; and above all, the assumption that νη- is a relic generally supplanted by ού is refuted by the clear absence of ού from composition. If νη- remained because the contours of the negative had been obscured by contraction, but the certainly much larger group with preconsonantal **ne-* assumed ού-, it is hard to see why there is in Greek not a single bahuvrīhi or synthetic showing ού-. The whole explanation is based on disregard of functional aspects.

Correspondences like νήκεστος : ἀνήκεστος : ἀνάκεστος, νηλεής : ἀνηλεής : ἀνελεής, νώνυμος : ἀνώνυμος²⁵ show traces of the IE allophones 2a and 2b; the leveling into types has here not led to the complete elimination of the non-syllabic grade, but to its limited coexistence with the syllabic variant. The origin of the lengthened vowel is obscure. The vocalic initials of the second elements presuppose an original laryngeal, and the largely 'prothetic' nature of these initials (e.g. ἀνήνωρ : ἀνῆρ, Ved. *nā*; νώνυμος : ὄνομα, Skt. *nāma*) makes us posit a primarily original **H-*. Whether the long vowel has to do with a primitive laryngeal or is due to a rhythmic principle in Greek, it is clear that it contains no

²⁴ E.g. Brugmann, KVG 310 (1933); Hirt 4.50.

²⁵ A useful though incomplete list by Gray, Lg. 1.119 f.

part of the first element; there are many cases which illustrate this, e.g. *δυσήκεστος*, *εὐώνυμος*, *ήμεύεις* (*νήνεμος*), *ήνυστρον* (*ἀνήνυστος*), *ήνορέα* (*ἀνήνωρ*). This theory also accounts for words like Gk. *νήπιος* (*ρεπος* < **Zmeq^uos*), where the original digamma has impeded satisfactory explanation. **νρεπιος* was a case like **νροθος*, the *ρ* being lost at an early stage; but contrary to *νόθος*, the word retained its compound character and joined the category of *νήκεστος*.

The grade *N-ñ-N*, /*ñ-N* was shortened in certain IE dialects (e.g. Germanic) and thus became indistinguishable from *N-η-N*. Where **ñ-* remained (Indic, Greek, Italic), the type was often doomed to isolation and disappearance through phonetic developments which separated it from **η-*. Some traces are distinguishable in Vedic meter: *āsat* < **NHηHsnt*, *ādevas*, *ārūpitas*. In Greek there are about a dozen words exhibiting *νη-* from **ñ-*, and in all cases but one there are analogical parallels with *ά-* (e.g. *νηκερδής*, *νήποτος*, *νηπευθής*; *νη(ρ)ις*). If the second *N* was a resonant, **ñ-* in Greek could derive its length only from a voiced laryngeal (except before *ξ* where the voiced laryngeal remained and the combination gave *ζ*, hence *ά-ζυξ*); the single case found supports this: *νη(ρ)ις*, Hitt. *hu-u-i-tar* (cf. IHL 50 f.). Before a non-syllabic resonant in Greek, a voiceless laryngeal remained after a syllabic as preaspiration of the resonant; it is seen in compounds like Hom. *ἀννέφελος*, *ἀνέφελος* < **ά-ην-*; *άλλοφος*.²⁶ Finally, if *an* be admitted as an Italic reflex of **ñ-*, Oscan-Umbrian *an-* may be considered a generalization of the results of *ηη-* and *ñ-*.

²⁶ Between non-syllabics the laryngeal was lost here as elsewhere (*νόθος*). Cases where *νέφος* 'makes position' originate in initial or postsyllabic position; the other variant is seen in *πολέμιοι νέφος* Il. 17.243.

CERTAIN OLD NORSE SECONDARY FORMATIONS

ALBERT MOREY STURTEVANT

University of Kansas

1. PRIMARY AND SECONDARY *lð* AND *rð*. The primary type of these two consonant clusters represents the PGmc. status when the two consonants were not due to the contraction of an intermediate vowel, whereas the secondary was the result of this contraction. In the primary type¹ both Gothic and Old Norse show that *l* exerted an occlusive force on the following *ð*; cf. PGmc. **stlē-stalðē* > **stlē-stal-ð* > Goth. *stai-stal-d* [not **stai-stal-þ*] 'acquired'; PGmc. **halðē* > PN **hal-ð* > **hal-d* [not **hal-þ*] > ON *hall* 'hold'. Only Gothic shows that *r* exerted an occlusive force upon *ð*; cf. Goth. *waur-d* [not **waur-þ*] 'word', but ON *or-ð*. For Old Norse we may then postulate the primary status as *ld* : *rð*.

In the later secondary type, due to the loss of an intervening vowel, the primary status was preserved in that *ld* replaced an earlier *lð*, while *rð* remained unaltered: cf. **taliðō* > *talða* > *talda* 'spoke', but **spuriðō* > *spurða* 'asked', never **spurda*; *talda* corresponds to *vil-da* < **wil-ðō* 'wished', and *spurða* to *orð*. It is clear, then, that in the secondary type *r* never exerted an occlusive force upon *ð*. Further evidence to this effect is that wherever the cluster *rd* (< **rdd*) arose through contraction, it developed secondarily to *rð*; cf. **herðiðō* > **hirðða* > *hirdda* > *hirda* > *hirða* 'tended', **werðiðō* > **wirðða* > *wirdda* > *virða* > *virða* 'valued'.

The *ð* in the forms *hirða* and *virða* is of later origin than the *ð* in *spurða*. In the secondary type *hirda* (> *hirða*), the cluster *rd* reverted to *rð*, the normal secondary type (as in *spurða*), which was also in accord with the primary type (*orð*). A parallel reversal to the primary type occurred in *lð* > *ld* (cf. *talða* > *talda*, on a level with *vil-da*). Later, another analogical form, *hirta*, was created beside *hirda* and *hirða*. Noreen explains *hirta* as derived from **hirð-ta*, with the *t*-suffix borrowed from the type *ræn-ta* 'robbed' in order to differentiate the tense forms pres. *hirða* and pret. *hirða*.² Though it is obvious that the *t*-suffix was borrowed from the type *ræn-ta*, Noreen's assumption that this borrowing had the purpose of differentiating the tense forms is far from convincing, in view of the types *senda* 'send' : *senda* 'sent', *skipta* 'shift' : *skipta* 'shifted'. It seems much more plausible to assume that the form *hirta* was derived not from a hypothetical form **hirð-ta* (as Noreen assumes) but from the phonetically correct form *hirda*. The voiceless stop *t* < *þ*, which was the result of assimilation with the preceding voiceless consonant, was never substituted for the voiced spirant *ð* but only for the corresponding voiced stop *d*; cf. the borrowing of *t* for *d* from the type *rænta* in *væna* 'to expect' : *vænda* > *vænta*, *þynna* (with original *nn*)

¹ The corresponding WGmc. forms of the primary type are omitted, since in West Germanic original **ð* suffered a spontaneous shift to *d*, which has no bearing upon the occlusive force of *l* and *r*.

² Cf. Noreen, *Aisl. Gram.*⁴ §238 Anm. 2: 'Spätere formen wie *hirta* (aus *hirð-ta*) haben zur verdeutlichung des tempus nach der analogie anderer verba mit lautgesetzlichem *t* (wie *ræna* rauben, pret. *rænta*) *t* angenommen.'

'to make thin': *þynda* > *þynta*. A proof of this hypothesis is that the *rt*-form appeared only when the *rd*-form was preserved beside the *rð*-form; thus *hirðā* : *hirða* > *hirta*, but never *virðā* > **virta* because the form *virta* was preserved only in the oldest MSS,³ and had disappeared by the time that *d* was supplanted by *t*. The *t* in *hirta*, then, could have been only indirectly borrowed from the type *rænta*, i.e. from the intermediate type *vænda* > *vænta* (cf. *hirða* > *hirta*), containing a stop *d* (< *ð) as a result of the occlusive force of the preceding nasal. The cluster *rd* in the type *hirða* then suffered two phonetic changes: either it reverted to the primary cluster *rð* (*hirða* > *hirðā*) or the *d* was supplanted by *t* (*hirða* > *hirta*) after the pattern of the type *vænda* > *vænta*. The retention of the anomalous form *hirta* may have been favored by the preterit type *skorta* (< **skortla* < **skort-þa* < **skort-ðā*) 'lacked'.

2. *Tega* 'TO SHOW'. This verb is not recorded in the preterit tense, but the past participle *tegat*, as well as the present tense of the mediopassive,⁴ shows that it belongs to the weak *ön*-class. It obviously represents a secondary formation based upon the stem **tig-* (with a low-grade vowel of the first ablaut series) of the strong verb **tīhan* (Goth. *teihan*) > *tjá* 'to point out, show', preserved in the adjective *tig-enn* 'pointed out, famous, distinguished' and in the derivative formations *tig-n* 'exalted rank, honor', *tig-na* 'to exalt, to honor'. Hence we should expect a form **tiga* instead of *tega*. All the *ön*-verbs of the type **tiga* retain the vowel *i*; only a very few show *a*-umlaut of *i* to *e*, and even these retain the form with original *i* as a doublet (cf. *dvīna*, *dvīn-enn* : *dvīna*, *dvena* 'to dwindle'; **svīna* [OHG *suīnan*], *svīn-enn* : *svīna*, *svena* 'to fade away, subside'). So far as I have discovered, the original vowel *i* was never completely discarded for *e* except in the case of **tiga* > *tega*.

If the vowel *e* in *tega* were due to *a*-umlaut, we should expect to find the form **tiga* persisting also. It is therefore plausible to assume a formal association between **tiga* and some other verb, whereby the original *i* was replaced by *e*. The contracted strong form of the infinitive *tjá* (< **tīhan*) led to the contracted weak form *tjá* (*tjáðā*, *tjáðr*) in the singular paradigm of the present indicative. When the new weak infinitive form **tiga* was created, there developed the relation *tjá* : **tiga*. Now the verb *mega* 'to be able' already exhibited the relation *má* (< **maχ* < **ma3*) : *mega* for the corresponding parts of the verb. The existing pair *má* : *mega* could have then led to *tjá* : *tega* (displacing the original **tiga*); the forms in both pairs correspond except for the initial consonant (cf. *sá* 'to sow' : *sera*, hence *slá* 'to strike' : *sló* > *slera*). This seems all the more plausible since **tiga* is the only *ön*-verb with the low-grade vowel of the first ablaut series in which the stem ended in *-g*. Note that the relation *má* : *megom* led to a similar relation in the paradigm of the synonymous verb *kná* (cf. OE *cnāwe*), namely *kná* : *knegom* (for **knōm* < **knāwum*). The same relation is preserved in *tjá* : *tegom*, where *tjá* represents the contracted form of the strong verb and *tegom* represents the new weak verb *tega*.

3. THE CONSONANTAL DECLENSION OF *ær* 'EWE'. The feminine substantive *ær*,

³ Cf. Heusler, *Aisl. Elementarbuch*³ §188; Ludvig Larsson, *Ordförrädet i de äldsta isländska handskrifterna* 368* (Lund, 1891).

⁴ Cf. *herr tegask verja land*, þormóðr Kolbrúnarskáld 2.18.

denoting a female animal, belongs to the consonantal inflection, parallel to *kýr* 'cow' and *sýr* 'sow'. Originally *ær* may have belonged to the *i*-stems (cf. Skt. *āvi-*, Gk. *δῆς*, Lat. *ovis*; OHG *ou*, plur. *auui euui*, OS *euui*), but may have passed over into the *jō*-stems (cf. OE *ewe eowu*). By assuming this development, we can perhaps explain why *ær* in historical times was transferred to the consonantal declension. If the earlier *jō*-declension had been preserved, the word would have had the following declensional forms: sing. *ær* (< **awīR*), **eyjar*, **eyju*, **ey*, plur. **eyj-ar*, *-a*, *-um*, *-ar*, parallel to *mær* 'maiden' (< **mawīR*), *meyjar*, *meyju*, *mey*, *meyj-ar*, *-a*, *-um*, *-ar*. All the declensional forms of *ær* except the nom. sing. would have been homonymous with those of *ey* 'island' (< **awī* < **awjō* < **a3wjō*). The formal identity between *ey* 'island' and **ey-* 'ewe' was removed when the substantive *ær* passed over into the consonantal declension; but this differentiation was probably only an incidental factor contributing to the shift of declension.

The point of departure for the shift was undoubtedly the retention of the *r*-ending in the nom. sing., which brought the form *ær*, with *R*-umlaut of the long radical vowel due to contraction (**awīR* > **āR* > *ær*), into direct contact with the consonantal declension of *kýr* and *sýr*, which also have *R*-umlaut of the long radical vowel (**kūR* > *kýr*, **sūR* > *sýr*) and which also denote female animals. The nom. sing. ending *-r* in *kýr* and *sýr* was never leveled out according to the oblique forms, as it was in the *jō*-stems *mær* (*mey*) and *þír* : *þýr* (*þý*) 'servant', which denote human females.⁵ The original *jō*-stem *ær* then discarded the pattern of *mær* and passed over into the pattern of the animal names *kýr* and *sýr*; cf. sing. *ær*, *ær*, *á*, *á*, plur. *ær*, *á*, *ám*, *ær*; *kýr*, *kýr*, *kú*, *kú*, plur. *kýr*, *kúa*, *kúm*, *kýr*. In this way the identity of the original *ey*-forms with those of *ey* 'island' was removed.

4. THE PREFIX **ga-* > *g-*. The preservation and loss of the unaccented prefix **ga-* was investigated in detail by Elis Wadstein, *Nordische Bildungen mit dem Präfix ga-*, IF 5.1-32 (1895). Wadstein's results have been to a large extent accepted as valid, but there is much in them that is uncertain. He rightly assumes that the preservation of the *g*-suffix (< **ga-*) was favored by the already existing pattern of primary initial *g* plus a consonant, resulting in the initial consonant clusters *gl*, *gn*, *gr*; thus *gladr* 'glad' (hence **ga-līkr* > *glīkr* 'like'), *gnaga* 'to gnaw' (hence **ga-nógr* > *gnógr* 'enough'), *grafa* 'to bury' (hence **ga-ranne* > *granne* 'neighbor'). But Wadstein made no attempt to explain why the *g*-prefix under these same phonetic conditions was often lost, nor did he take into consideration that an ON form may not have contained the prefix **ga-* which appears in cognate

⁵ The primary form *mær* was the regular form in both poetry and prose; the secondary form *mey* appeared only in late prose. On the other hand, the primary form *þír* (> *þýr*) was retained only occasionally in poetry and the þulur; elsewhere it was completely discarded for the secondary form *þý* in both poetry and prose. The retention of the original form *mær* (with the case ending *-r*) during the whole ON literary period, contrary to the pattern of *þír*, may be explained as due to the formal influence of *ær*, which contained the original **R*-ending of the nom. sing. case of the *jō*-declension (cf. **awīR* > *ær*, **mawīR* > *mær*). This formal association checked the leveling tendency, which was unrestricted in the case of *þír*.

forms before *l*, *n*, or *r* in the other old Gmc. dialects. It is these factors to which I should here like to call attention. All the following examples have reference to *g* plus *h*, *l*, *n*, *r*, in which the *g*-prefix is either preserved or lost when we should expect it to be uniformly preserved.

(a) There is no evidence that the prefix *g*- in the forms *glíkr* 'like', *gnógr* 'enough' disappeared in the forms *líkr*, *nógr*; the latter may represent independent formations (cf. OS *gi-lík* : OFris. *lik*). There are no clear examples in which the *g*-prefix has been both preserved and lost in the same word.

(b) When forms with the *g*-prefix in the other old Gmc. dialects are cognate with an ON form without the prefix, the latter may represent loss of an earlier *g*- or an independent formation that never had the prefix. The choice must be settled entirely by the semantic implications of the prefix **ga*-. Where the sociative connotation was present in a cognate word, it may not have been present in the ON word. Compare ON *g-ranne* 'neighbor' with *rúne* 'secret adviser': *g-ranne* is cognate with Goth. *ga-razna*, *rúne* with OE *ge-rūna* and OHG *ga-rūno*. The prefix **ga*- in both the Gothic and the WGmc. cognates obviously had a sociative implication. In ON *g-ranne*, the *g*-prefix still preserved this implication: 'one who dwells in the same house (ON *rann* : Goth. *razn*)'. In ON *rúne*, on the other hand, there is no convincing evidence (in spite of the WGmc. cognates) that the **ga*-prefix ever was present. The word *rúne* (< PN **rūnē*) could well be the semantic equivalent of *ráð-gjafi* 'one who gives counsel, a counsellor', but with a connotation of 'secret, private' (cf. *rún* 'a secret, mystical sign'), hence 'an intimate friend'.

(c) When it is certain that the **ga*-prefix had a sociative implication in ON but nevertheless disappeared, its loss may have been due to secondary ON conditions. Thus, in ON *lande* (OE *ge-landa* : OHG *gi-lanto*) 'one who is from the SAME land as another person, a compatriot' and ON *nafne* (OE *ge-namna* : OHG *gi-namno*) 'one who has the SAME name as another person, a namesake', the loss of the *g*-prefix (present in the WGmc. cognates) can be explained as due to the influence of compound forms with the prefix *sam*- 'same', *sam-lande* and *sam-nafne*, which rendered the *g*-prefix superfluous. That is, the simple forms *lande* and *nafne* may have been abstracted from the synonymous *sam*-formations, and in this way the original *g*-prefix may have been discarded.

(d) When a verb without the *g*-prefix denotes a perfective act by itself, it is not certain that this prefix ever existed in the ON verb, even when it is present in cognates in the other Gmc. dialects. Thus, the verb *leiða* 'to lead' denotes a perfective act, 'to get someone to go or come (*liða*), kommen lassen' (OHG *ge-leiten*, NHG *be-gleiten*), and therefore may never have had the **ga*-prefix. Hence it may be better equated with the WGmc. forms which lack this prefix: OHG *leiten*, OS *lēdian*, OE *lēdan*.

(e) Another factor to be considered is a possible shift in the verbal government. For instance, the verb *heita* means 'to call, to name'; but with the **ga*-prefix it came to denote a perfective act, 'to promise, verheissen' (cf. Goth. *ga-haitan*). This shift in meaning involved a corresponding shift in the verbal government: *heita* takes an accusative object, whereas **gaheita* takes a dative object denoting

the person to whom something is promised. The syntactical shift indicated the perfective act without the use of the *g*-prefix (**ga-heita* > **geita*, cf. *eisa* : *g-eisa* 'to rush'), and hence favored its loss.⁶

5. CONTRACTION IN THE PAST PARTICIPLE OF WEAK *jan*-VERBS WITH SHORT STEM SYLLABLES. Both the contracted and the uncontracted forms occur in the past participle of weak *jan*-verbs with short stem syllables, except when the dental suffix **ð* in the preterit forms follows a dental (*ð*, *t*) of the stem syllable, resulting (through syncope of the intermediate vowel *i*) in the geminates *-dd-* and *-tt-*. So far as the question of contraction is concerned, we may designate the two types of verbs as non-dental and dental. Examples of the non-dental type: *ver-ja* 'to defend', ppl. *var-i-ðr* > *var-ðr*, pret. *var-ðā*; *dyl-ja* 'to conceal', ppl. *dul-i-ðr* > *dul-ðr* > *dul-dr*, pret. *dul-ðā* > *dul-da*; *vek-ja* 'to awaken', ppl. *vak-i-ðr* > *vak-þr* > *vak-tr*, pret. *vak-þā* > *vak-ta*; *glep-ja* 'to deceive', ppl. *glap-i-ðr* > *glap-þr* > *glap-tr*, pret. *glap-þā* > *glap-ta*. Examples of the dental type: *kveð-ja* 'to speak', ppl. *kvað-dr* (never **kvað-i-ðr*), pret. *kvað-da*; *flyt-ja* 'to move', ppl. *flut-tr* (never **flut-i-ðr*), pret. *flut-ta*. Here the uncontracted forms were leveled out in favor of the contracted forms already in the prehistoric era; the uncontracted forms do not occur even in the oldest MSS. The question arises, what caused this chronological discrepancy in contraction between the dental and the non-dental type of past participle. So far as I know, this question has never been answered.

The fundamental distinction between the two types of past participle is the fact that in the dental type the contraction resulted in the geminates *-dd-* and *-tt-*, whereas in the non-dental type no geminates could develop through contraction. In both types the dental suffix *ð* of the contracted participle underwent the same phonetic changes as in the contracted preterit form (cf. pret. *vak-þā* > *vak-ta* : past part. *vak-þr* > *vak-tr*; pret. *kvað-da* : past part. *kvað-dr*; pret. *flut-ta* : past part. *flut-tr*); these two forms were therefore always on a level, so far as the dental suffix *ð* is concerned. The preterit forms *kvadda* and *flutta*, however, were anomalous in that (contrary to the non-dental type) the contraction resulted in the geminates *-dd-*, *-tt-*. The uncontracted forms of the corresponding past participles, **kvað-i-ðr* and **flut-i-ðr*, were first discarded (during the prehistoric era) for *kva-dd-r* and *flu-tt-r* because only through contraction could the geminates *-dd-* and *-tt-* of the preterit be preserved in the participle. Later, during the historical period, the non-dental class followed the pattern of the dental class, but with preservation of the original final consonant of the stem syllable (cf. *vak-i-ðr* > *vak-þr* > *vak-tr* : pret. *vak-þā* > *vak-ta*, but **kvað-i-ðr* > *kva-dd-r* : pret. *kva-dd-a*). The dental type **flut-i-ðr* > *flut-tr* likewise preserved the original final consonant of the stem syllable (cf. *flyt-ja* : pret. *flut-ta*), but **flut-i-ðr* > *flu-tt-r* obviously followed the pattern of **kvað-i-ðr* > *kva-dd-r* because of the geminates *-tt-* and *-dd-* resulting from contraction.

6. THE UNCONTRACTED MIDDLE SYLLABLE IN THE MASCULINE *a*-DECLENSION OF PROPER NAMES. In dissyllabic appellatives of the masculine *a*-declension, contrac-

⁶ Cf. the substitution of the particles *of* and *um* for the **ga*-prefix in connection with both verbs and substantives. This substitution, however, does not explain the loss of the prefix.

tion of the middle syllable took place when a vowel of the ending was added (i.e. in the dative singular and throughout the plural paradigm); cf. *hamar-r* 'hammer', *hamar-s*, *hamr-i*, *hamar*, plur. *hamr-ar*, *-a*, *hōmr-um*, *hamr-ar*; *aplan-n* 'evening', *aplan-s*, *aptn-i*, *aplan*, plur. *aptn-ar*, *-a*, *ōptn-um*, *aptn-ar*. Proper names are regularly confined to the singular; in the dative, those with the two suffixes *-ar* and *-an* never show contraction; cf. *Gunn-ar-i*, *þór-ar-i*, *Herj-an-i*.

This discrepancy may be partly explained by supposing that the contracted form existed at an earlier period (if the vowel *a* of the suffix represents an original short vowel) but that since proper names occur only in the singular, this form was later leveled out in favor of the uncontracted forms in the rest of the paradigm (thus **Gunþ-harjær* > *Gunnar-r*, *Gunnar-s*, **Gunn-ri* > *Gunn-ar-i*, *Gunnar*). In the appellatives, on the other hand, the preservation of the contracted form in the dative singular case was favored by the corresponding contraction in the entire plural paradigm. Note that the suffix *-ar* in compound proper names is not identical in origin with the suffix *-ar* in appellatives; hence where the short vowel *a* in the suffix *-ar* of proper names goes back to an earlier long **ā*, the contraction never occurred: thus *þór-ar-r* (**þór-ār-ak* < **þór-gār-ak* < **þór-gair-ak*), dative *þór-ar-i* (< **þór-ār-i*); cf. *þjóð-mar-r* (< **mār-ir*), dative *þjóð-mar-i* (< **mār-i*), and the suffix **-ār-ē* > *-ar-i* in **harp-ār-ē* > *harp-ar-i* etc. The late ecclesiastic name *Pét-arr* (dat. *Pét-ar-i*) was undoubtedly modeled on the pattern of the native type *Gunn-arr*, *þór-arr* (cf. *Pét-urr* : *Giz-urr*). The foreign appellative *bik-ar-r* 'basin' (< Lat. *bicār-ium*) then followed the pattern of the foreign type established for proper names with the suffix *-ar* (dat. *bik-ar-i* : *Pét-ar-i*).

On the other hand, the suffix *-an* in the proper name *Herj-an-n* is identical in origin with the *an*-suffix of the appellative *apt-an-n*. Although the name *Herj-an-n* represents an appellative *herj-an-n* 'chieftain', it is not recorded as an appellative but only as a name for Odin. Since the proper name has no plural paradigm, the dative singular form *Herj-an-i* (without contraction) can be explained in no other way than through leveling in the singular. The same principle applies to the appellative *þjóð-an-n* (cf. Goth. *þiud-an-s*, OE *þeod-en*) 'the head of a people' (*þjóð*, Goth. *þiud-a*), hence 'chieftain, king'. The uncontracted dative singular *þjóð-an-i* is parallel to *Herj-an-i* and at variance with *apt-ni*; it is significant that the appellative *þjóð-an-n*, like the proper name *Herj-an-n*, never occurs in the plural.

7. THE PROPER NAME *Half-dan(r)*. In compound proper names the original final consonant cluster *-nr* (< **-nr*) was, through lack of accentuation, regularly assimilated to *-nn* and later simplified to *-n*; cf. **Auð-un-r* > *Auðun(n)*, **Hák-un-r* > *Hákun(n)*. But in the name *Half-danr*, the final cluster *-nr* remained exempt from this assimilation; it was a subsequent loss of the ending *-r* that produced the secondary form *Half-dan*. The discrepancy between type 1 (*Auð-un[n]*) and type 2 (*Half-dan[r]*) has never been explained, perhaps because one possibility has not been considered, viz. the difference in accentuation between the two types. In type 1 the element *-unn* obviously represents the simplex *vinr* 'friend' (cf. *Auð-unn* : OE *Éad-wine*, *Hák-unn* : OHG *Hanc-win*); in unaccented position, **-winr* > **-wunn* > *-unn*. In type 2, on the other hand, it seems plausible to assume that the main stress was originally on the final element (*Half-dánr*),

i.e. 'half Dáne' in contradistinction to a half Swéde or the like; cf. *half-vinr*⁷ 'half-friend' instead of **half-un-n* parallel to *Auð-un-n*. This assumption also explains the retention of the ending *-r* in the cluster *-nr*; for after a short vowel of an accented syllable the assimilation of *-nr* (< **-næ*) to *-nn* never took place (cf. *svanr* 'swan' : *-danr* 'Dane').

To account for the subsequent loss of the final *-r* we must assume that the main stress was shifted from the second element of the compound to the first, in accord with a general Gmc. tendency.⁸ The cluster *-nr* in the unaccented element *-danr* was then simply carried over from the accented element *-danr* through association with the simplex, and so remained exempt from the assimilation to *nn* which took place in the unaccented syllables of type 1. The lack of accent resulted only in the loss of the final *-r*. The assumption that the form *Half-dan* was derived not from *Half-danr* but from **Half-dann* (< **Half-danæ*), on a level with *Auð-un* < *Auð-unn* (< **Auð-unæ*), leaves unexplained the existing form *Half-danr*.⁹ The discrepancy can be explained by the shift of accentuation, as shown above.

8. THE REDUPLICATED PRETERITS *séra* AND *snóra*. The regular preterits of *sá* 'to sow' and *snúa* 'to twist, to turn' are *sera* (< **sezō* = Goth. *saisō*) and *snóra* for *snera* (< **snezō*), both with short radical vowels. The vowel lengthening in the late secondary forms is explained by Noreen as borrowed from the long radical vowels *á* and *ú* of the present system and the past participle (cf. *sá*, *sáenn*; *snúa*, *snúenn*).¹⁰ This hypothesis seems a priori untenable, since we have no parallels for the assumed lengthening, in which the original short vowels and the long ones that served as models were of different qualities (*e* : *á*, *e* : *ú*). Analogical vowel lengthening occurs only as a change in the quantity of the same vowel, as when *ǣ* > *ā* in *nam* > *nām* : *nāmun* 'took' after the pattern of *át* : *átun* 'ate'. Against Noreen's hypothesis is also the fact that the vowel lengthening was restricted to the reduplicating type with initial *s-* (*sera*, *snóra*), and was not extended to the type based upon initial *r*; we never find *réra* (*róa* 'to row') : *g-réra* (*gróa* 'to grow').

For the discussion of this question we divide the verbs with reduplicated preterit into two classes: those with initial *s-*, and those with initial *r*. Since *g-núa* 'to rub' followed the pattern of *snúa*, and *gróa* followed the pattern of *róa*,¹¹ it is enough to discuss the verbs *sá* and *róa* as representing these two classes.

The infinitive *sá* (< **sā-a*, cf. OHG *sā-en*) represents a contracted monosyllabic form with long radical vowel *ā*, on a level with the contracted type *tjá* 'to show' < *téa* < **tīhan* (Goth. *teihan*), *ljá* 'to lend' < *léa* < **līhwan* (Goth. *leihwan*). When the two latter verbs passed over into the weak conjugation, a new vowel pattern

⁷ See Fritzner, *Ordbog over det gamle norske sprog*² 707^a.

⁸ Cf. *þu-ríðr* > *þý-ríðr*, *Guð-ríðr* > *Gv-ríðr*.

⁹ Cf. Noreen §285 Anm. 2.

¹⁰ Cf. Noreen §126 Anm. 2: 'Die selt. prät.-formen *séra*, *snóra* statt *sera* (got. *saisō*) säete, *snóra* wandte haben wol die länge aus dem inf. *sá*, *snúa*, resp. dem präs. und part. entlehnt.'

¹¹ On the relation of these reduplicating verbs to one another seen Scandinavian studies 23.64-5 (1951).

was created for the present and preterit forms, viz. $\acute{a} : \acute{e}$; cf. *tjá*, *léða*, *léðr*,¹² *ljá*, *léða*, *léðr*. Therefore it seems plausible to assume that the short vowel *e* in the form *sera* was lengthened according to the proportion $tjá : léða = ljá : léða = sá : sera > séra$. The infinitives *tjá* and *ljá* represent contracted strong verbs with long vowel \bar{a} , as does *sá*; the preterits *léða* and *léðr* represent secondary weak formations, as does *sera*. This originally represented a strong form; but on account of the final *-a* it passed over into the weak conjugation, parallel to *téð-a* and *léð-a*. Since the radical vowel *e* alternated with ϕ in *snera* : *snøra*, the lengthening could easily be extended from *e* to ϕ , yielding *snøra*. That the forms **séra* (for *søra*) and **snéra* (for *snera*) did not occur may be due to the sporadic character of the vowel lengthening.

As for the class with initial *r-*, since the infinitive form *róa* is dissyllabic and contains the radical vowel \acute{o} , it could not be affected by the contracted type *tjá* and *ljá*, from which the secondary pattern $\acute{a} : \acute{e}$ developed.

¹² The forms *léða* and *léðr* were based upon the infinitive form *léa*, *tjáða* and *tjáðr* upon the infinitive form *tjá*.

TOWARD THE LOGICAL DESCRIPTION OF LANGUAGES IN THEIR PHONEMIC ASPECT

E. COLIN CHERRY
University of London

MORRIS HALLE
*Massachusetts Institute
of Technology*

ROMAN JAKOBSON
Harvard University

Distinctive features occur in lumps or bundles, each one of which we call a phoneme. The speaker has been trained to make sound-producing movements in such a way that the phoneme-features will be present in the sound-waves, and he has been trained to respond only to these features and to ignore the rest of the gross acoustic mass that reaches his ears. LEONARD BLOOMFIELD (1933)

The number of different phonemes in a language is a small submultiple of the number of forms. LEONARD BLOOMFIELD (1926)

The logical demand that a science speak in quantitative terms is met by linguistics because it speaks in terms of phonemes. LEONARD BLOOMFIELD (1927)

1. INTRODUCTORY

This paper, an attempt to contribute to a logical description of the phonemic structure of a language, employs some of the elementary concepts of statistical communication theory.¹ Particular illustration is provided by a statistical analysis of colloquial Russian;² the material studied was the Russian urban conversations recorded by Peškovskij, comprising in the latter's phonetic transcription ten thousand sounds.³

In analyzing Russian or any other language, we must ascertain what and how many DISTINCTIVE FEATURES are needed to differentiate the meaningful units of its code, i.e. the smallest meaningful units, termed morphemes, and their combinations into words. Words are the maximum units that are expected to be entirely provided by the code. We must determine the minimum set of such features that the listener needs in order to recognize and distinguish all except homonymic morphemes, without help from context or situation. Once this set is determined, all other phonetic differences among morphemes or words of the given language can be shown to be predictable and therefore redundant.⁴

If we compare, for example, the Russian words (1) [bit] 'way of life', (2) [b, it]

¹ See, in particular, C. E. Shannon and W. Weaver, *The mathematical theory of communication* (Urbana, 1949); D. M. Mackay, In search of basic symbols, *Cybernetics: Transactions of the eighth Conference* (New York, 1952); id., The nomenclature of information theory, *ibid.*

² This analysis was made as part of the research on contemporary Russian conducted by the Department of Slavic Languages and Literatures at Harvard University under a grant from the Rockefeller Foundation. The first volume of the description of contemporary Russian based on this research, dealing with the various aspects of Russian speech sounds, is now being prepared for publication. Grateful acknowledgment of help is also made to the Signal Corps, Air Materiel Command, Office of Naval Research, and the sponsors of the Fulbright program.

³ A. Peškovskij, Desjat' tysjač zvukov russkogo jazyka, *Sbornik statej* 167-91 (Leningrad, 1952).

⁴ For further information on distinctive features and their acoustic and articulatory correlates, see R. Jakobson, C. G. M. Fant, and M. Halle, *Preliminaries to speech analysis*, 2d printing (MIT Acoustics Laboratory, Technical report No. 13, 1952).

'beaten', (3) [bit,] 'be', and (4) [b,it,] 'beat',⁵ we observe that words (1) and (2), or words (3) and (4), differ from each other in two respects: [ɪ] is farther forward than [i] (i.e. has a higher second formant), and [i] is farther forward than [i]; while [b,] is distinguished by its palatalization from [b]: it is produced with a flattening of the mouth cavity and a simultaneous widening of the pharyngeal channel which results in an upward displacement of energy along the frequency axis. Words (1) and (3), or words (2) and (4), also differ from each other in two respects: [i] is closer than [ɪ], and [i] is closer than [ɪ]; while [t,] again differs from [t] in its palatalization.⁶

If, now, the presence as opposed to the absence of consonantal palatalization is viewed as the distinctive feature, this one feature (which we call sharp vs. plain) suffices to differentiate the four words in question. If, on the other hand, the distinctive function were to be assigned to the vowels, we would have to postulate two independent features, front vs. back and close vs. open. This superfluous multiplication of features is reason enough for rejecting the second solution. In Russian there is an additional reason for adopting the first solution, for in this language, in certain positions, the presence or absence of consonantal palatalization can by itself distinguish sequences otherwise identical, and must therefore in any case be considered an autonomous distinctive feature; compare [voʃt,] 'leader' with [koʃt] 'outlay', or [sɛl,t,] 'herring' with [k,ɛl,t] 'Kelt'. The distribution of advanced and retracted, on the other hand, as well as that of close and open varieties of vowels, is entirely predictable from the presence or absence of palatalized consonants before and after the vowel.

Proceeding consistently in this way, we find in the code of contemporary Standard Russian eleven distinctive features, grouped by superposition into forty-two phonemes.⁷ These eleven distinctive features suffice to differentiate all but homonymic morphemes and words in Russian.

We leave aside here sound features that perform other functions, namely CONFIGURATIONAL features that signal the division of the utterance into grammatical units of different degrees of complexity, and expressive (or more precisely PHYSIOGNOMIC) features that signal solely the emotional attitudes of the speaker. Examples of configurational features signaling the division of the sound chain into word units: [dɐv'ol,nij] /da v'ol,nij/ 'free besides' : [dav'ol,nij] /dav'ol,nij/ 'content'; [t,ɛ'n,it'am] /t,ɛn,i tam/ 'shadows are there' : [t,ɛn,it'am] /t,ɛn,i-t'am/ 'they are elsewhere'; [jix'i'də jix,i'də] /j'ix 'ida. jix'ida/ 'their Ida is

⁵ Cf. A. Isačenko, *Fonetika spisovnej ruštiny* 177, 182 (Bratislava, 1947).

⁶ We follow the IPA system of transcription, except in three respects: we use a comma after a letter to indicate palatalization; we place the accent mark immediately before the vowel letter; and we render the strident stop by the same letter as the corresponding constrictive with the addition of a circumflex.

⁷ There are two competing varieties of contemporary standard Russian. The more conservative is codified especially in *Tolkovyj slovar' russkogo jazyka*, ed. by D. Ušakov (Moscow, 1935-40); the other is advocated in particular by S. Obnorskij, and is presented in *Slovar' russkogo jazyka*, ed. by S. Ožegov (Moscow, 1949). In general we accept Ušakov's norms; but in order to include all the phonemic discriminations possible in standard Russian, we add to his traditional repertory of phonemes a new phoneme /g,/ as distinguished from /g/. Such new gerund formations as /b,ir,ig,'a/ 'taking care', distinct from /b,ir,ig'a/ 'banks', are admitted into standard Russian by Obnorskij and his followers.

malicious'. Physiognomic features are illustrated in the different ways of pronouncing the word for 'yes' (simply [d'a] when unemphatic) according to the degree and kind of emphasis. These features convey subsidiary information similar to what is carried by such graphic equivalents of configurational features as spaces or punctuation marks, and such equivalents of physiognomic features as underlining or italicizing. The REDUNDANT features, on the other hand, operate in conjunction with the distinctive features, thereby facilitating the selective process on the part of the listener and lessening the burden on his attention.

For our computations, the text was split up into phoneme sequences consisting of two successive vowels and the consonants (if any) between them. In this way each vowel appears twice in our corpus, once as the initial and once as the final phoneme of a sequence. We chose these sequences 'from vowel to vowel' because phonemic conditioning is confined, in Russian, to consonantal clusters and to combinations of a vowel with preceding or following consonants; there is no apparent influence on consonants following a given vowel by those preceding it or vice versa. The compulsory syntactic pause (both initial and final) was denoted by a period and equated with a vowel.

Three sets of counts are of interest: (A) those that regard both the word boundaries (symbolized by a space) and the junctures between the immediate constituents of compound words⁸ (symbolized by a hyphen); (B) those that regard only the word boundaries; and (C) those that regard neither the word boundaries nor the junctures, but break up a sequence only at the points of compulsory pause. The three ways of dividing a text into elementary sequences are illustrated in the accompanying table, based on the following passage: *Vot, na tebe na obed. Pojděš' ...* /v'ot. n'a t,ib,e na-ab,et. pa-jd,oʃ./ 'Here, that's for your dinner. You'll go ...' The computations in this paper are made according to the first way of counting.

(A)	v'o	'otʃ	ʃn'a	'aʃ	ʃt,i	ib,e	'eʃ	ʃna	aʃ	ʃa	ab,e	'etʃ	ʃpa	aʃ	ʃjd,o	'oʃʃ
(B)	v'o	'otʃ	ʃn'a	'aʃ	ʃt,i	ib,e	'eʃ	ʃna	aa	ab,e	'etʃ	ʃpa	ajd,o	'oʃʃ		
(C)	v'o	'otʃ	ʃn'a	'at,i	ib,e	'ena	aa	ab,e	'etʃ	ʃpa	ajd,o	'oʃʃ				

2. THE FEATURE PATTERN AS A LOGICAL DESCRIPTION OF THE PHONEME

In the description that follows, language will be treated as a Markoff process.⁹ The phonemes will be considered uniquely identifiable; but their order, in the sequences that compose our sample, can be described only statistically.

⁸ Among Russian compound words we include all words with a non-initial root: words with more than one root, e.g. /adna-abr'aznij/ 'uniform'; words with prefixes, e.g. /za-astr,'it,/ 'to sharpen', /iz-vad,'it,/ 'to exhaust'; and words with preceding prepositions which are phonemically treated like prefixes, e.g. /za-akn'o/ 'behind the window', /iz-vad'i/ 'out of the water'.

⁹ Cf. Shannon and Weaver 102: 'A system which produces a sequence of symbols ... according to certain probabilities is called a *stochastic process*, and the special case of a stochastic process in which the probabilities depend on the previous events, is called a Markoff process or a Markoff chain.' In his *Essai d'une recherche statistique sur le texte du roman 'Eugène Onégin'*, illustrant la liaison des épreuves en chaîne, *Bulletin de l'Académie Impériale des Sciences de St. Pétersbourg*, Vol. 7 (1913), A. A. Markov studied the distribution of vowel and consonant LETTERS in a part of Puškin's famous poem and showed that the transitional probabilities between the letters were not those of a random sequence but rather depended on the preceding letter or letters.

For the task of identifying one particular phoneme out of the set employed by the language, the distinctive features may be regarded as questions to be answered yes or no. Thus one may ask, Is the phoneme vocalic?—yes or no; Is the phoneme consonantal?—yes or no; and so on through the entire list of features. For the language under consideration here, a total of eleven such questions is necessary to identify any one phoneme uniquely. Table A illustrates these questions answered yes (+) or no (-); a zero (0) means either. This suggests that the logic is three-valued, a point that will be taken up again later.

A simple illustration of such a logical description is provided by Fig. 1, which shows a set of eight 'objects' A, B, ... H, to be identified by yes (+) or no (-) answers. Thus the group is first split in two, and we begin by asking, Is the object that we want on the right side (+) or not (-)? Successive subdivisions eventually identify any object in a set. If there are N objects in the set, and if N happens to be a power of 2, the number of yes-or-no answers necessary to identify each of the objects in the set is $\log_2 N$. The complete identification of any object is then a chain of plus and minus signs; thus, the object G in Fig. 1 is identified by the chain (+ + -).

A	B	C	D	E	F	G	H
-	-	-	-	+	+	+	+
-	-	+	+	-	-	+	+
-	+	-	+	-	+	-	+

FIG. 1. THE LOGICAL IDENTIFICATION OF OBJECTS IN A SET OF EIGHT

Even when N is not a power of 2, the quantity $\log_2 N$ can still be used as a measure. In such cases the fractional result must not be taken to imply a fraction of a question; it means, rather, that the N members of the set will not all necessarily require the same number of answers for identification. The fraction results from averaging.

The quantity $\log_2 N$ is conventionally expressed in BITS; the name for this unit is derived from BINARY DIGIT (i.e. yes-or-no choice).

In Fig. 1 the successive subdivision has been consistently into two equal subgroups; this method results in identification by the smallest possible number of answers, and so in the shortest chain of plus and minus signs. Subdivision into unequal subgroups requires, on the average, more questions and answers.

Let us now apply this process to the list of forty-two Russian phonemes listed in Table A. But first consider a purely hypothetical description of any one phoneme out of the forty-two, as though these were not phonemes but merely objects without linguistic significance. If they were successively subdivided as in Fig. 1, the description of any one object would require $\log_2 42$ questions, on the average, or 5.38 bits per phoneme. In our analysis of language we are concerned, however, not only with questions of logic but also with matters of fact; hence the answers yes or no in Table A are provided for us by considerations of the natural process of speaking.

One might ask, Why cannot a type of feature pattern be invented which

employs only 5.38 questions per phoneme, on the average, in a manner analogous to the hypothetical case discussed? This could perhaps be done; but the distinctive features used at present (Table A) serve other purposes and are intimately related to the physical production of speech. They number eleven, implying an average of 5.62 extra questions per phoneme ($11 - 5.38$). This means that redundant or extra plus and minus signs are brought in. Nevertheless these features, as they have been proposed for earlier linguistic analyses, fit into the logical

	k	k	g	g	x	c	ʃ	ʒ	t	t	d	d	s	s	z	z	ʂ	n	n	p	p
VOCALIC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CONSONANTAL	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
COMPACT	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DIFFUSE	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
GRAVE	+	+	+	+	+	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+
NASAL	0	0	0	0	0	0	0	0	-	-	-	-	-	-	-	-	-	-	+	+	-
CONTINUANT	-	-	-	-	+	-	+	+	-	-	-	-	+	+	+	+	-	0	0	-	-
VOICED	-	-	+	+	0	0	-	+	-	-	+	+	-	-	+	+	-	-	-	-	-
SHARP	-	+	-	+	0	0	0	0	-	+	-	+	-	+	-	+	0	-	+	-	+
STRIDENT	0	0	0	0	0	0	0	0	-	-	-	-	0	0	0	0	+	0	0	0	0
STRESSED	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	b	b	f	f	v	v	m	m	ʲu	ʲo	ʲe	ʲi	ʲa	a	r	r	l	l	j
VOCALIC	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	-
CONSONANTAL	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	+	+
COMPACT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+	0	0
DIFFUSE	0	0	0	0	0	0	0	0	+	+	-	-	+	+	0	0	0	0	0
GRAVE	+	+	+	+	+	+	+	+	+	+	+	-	-	0	0	0	0	0	0
NASAL	-	-	-	-	-	-	+	+	0	0	0	0	0	0	0	0	0	0	0
CONTINUANT	-	-	+	+	+	+	-	-	0	0	0	0	0	0	0	0	-	+	+
VOICED	+	+	-	-	+	+	-	-	0	0	0	0	0	0	0	0	0	0	0
SHARP	-	+	-	+	-	+	-	+	0	0	0	0	0	0	0	0	-	+	+
STRIDENT	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
STRESSED	0	0	0	0	0	0	0	0	+	-	0	0	+	-	+	-	0	0	0

TABLE A. THE PHONEMES OF RUSSIAN
showing their distinctive feature patterns as represented by the answers
yes (+), no (-), either (0)

descriptive system, though apparently with some inefficiency. Can the efficiency of our empirical description be improved by simplification of Table A?

Table A shows the simplest possible description of the 42 phonemes in terms of the given eleven features. There are several points of difference between this table and Fig. 1. First, the successive questions have phonetic significance; they do not merely ask Right or Left? like those in Fig. 1. The answer to the first question (vocalic—yes or no?) does not split the 42 phonemes into two equal groups, but into 12 pluses and 30 minuses; Russian phonemes simply happen to have this characteristic. The second question (consonantal—yes or no?) again parts each of these groups into unequal subgroups; and so on.

Moreover, some of the questions in the list need not be answered at all for

particular phonemes, because the identification is complete without them. In Table A we use a zero to indicate 'either'—that is, either plus or minus. For example, the phoneme /t/ is represented by the chain $(- + - 0 - - - - 0)$. Each of the zeros can be replaced by either plus or minus without affecting the identification; in either case, the chain of symbols for /t/ remains unique. Since every zero may thus be regarded as either a plus or a minus, the total number of questions answered here is eleven per phoneme. This is a measure of the 'information' conveyed when the speaker selects any particular phoneme out of the 42, at least on the basis of the feature pattern here presented. But as we have seen, the true 'information' is rather to be expressed by an average of 5.38 questions (bits); the extra 5.62 bits represents the redundancy that would result from the replacement of the zeros by plus or minus signs. (It must be emphasized that our measure of 'information' has up to this point been based upon the assumption that all 42 phonemes have an equal probability of occurrence and that they are wholly independent units. Since language has, of course, a much more complex structure than this, our definition of 'information' will later have to be modified.)

The term 'redundancy' should not be taken to imply wastefulness; it is a property of speech, and in fact of every system of communication, which serves a most useful purpose. In particular, it helps the hearer to resolve uncertainties introduced by distortion of the signal or by disturbing noises. For example, the feature of nasality is marked 0 for all vowels. If these zeros were changed to pluses, the new symbols would not imply that a Russian speaker always nasalizes his vowels: normally he does not; but even if he did, the nasality would have no phonemic significance. In some cases a zero appears in a place where the substitution of plus or minus would imply an impossible articulation; but even here the point is that the phoneme is uniquely identified without this feature.

If the data given in Table A can be recast so as to eliminate the necessity of using the ambiguous symbol 0, then the number of questions needed to identify any one phoneme will, on the average, be reduced. That is, the description of phonemes in terms of features will be less redundant.

3. REMOVAL OF THE AMBIGUOUS ZERO SIGNS

One might suppose that by re-ordering the feature questions, it would be possible to remove all the zero signs in Table A, or at least to shift them to the end of every phoneme column so that they could be omitted (the phoneme being identified then by the chain of plus and minus signs only). It turns out, however, that this cannot be accomplished by any simple re-ordering.

The whole problem may be changed by regarding the table of signs (+, -, 0) as a code book for identifying the various phonemes. In this view there is no reason why the order of the feature questions should not be different for different phonemes. In fact, the order could change during the identification of a particular phoneme, at certain stages depending upon the answers to earlier questions. Thus a sequence of different code books would be required. Table B shows the result of such a recoding.

As an example, consider the identification of the phoneme /^lo/. The answers to the questions Vocalic? Consonantal? Compact? are respectively + - -,

which identifies the phoneme as belonging to the group /'u u 'o 'e 'i i/. This requires that a new code book be used for the subsequent questions. These, as we see from Table B, are asked in the order Diffuse? Grave? Stressed? The code books are known a priori and represent here the independent phoneme structure of Russian; they themselves contain the 'information' provided by the zeros in Table A.

This process of recoding may be regarded as a TRANSFORMATION. The number of signs (bits) required to identify any phoneme uniquely is now less than before

	k	k	g	g	x	c	ʃ	ʒ	'u	u	'o	'e	'i	i	'a	a	r	r	l	l	j
VOCALIC	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+	+	-
CONSONANTAL	+	+	+	+	+	+	+	+	-	-	-	-	-	-	-	-	+	+	+	+	-
COMPACT	+	+	+	+	+	+	+	+	-	-	-	-	-	-	+	+					
GRAVE	+	+	+	+	+	+	+	-													
CONTINUANT	-	-	-	-	-	+	-	+													
VOICED	-	-	+	+				-	+												
SHARP	-	+	-	+																	
	DIFFUSE								+	+	-	-	+	+							
	GRAVE								+	+	+	-	-	-							
	STRESSED								+	-			+	-	+	-					
									SHARP												
									CONTINUANT								-	+	-	+	
																	-	-	+	+	

	t	t	d	d	s	s	z	z	ʃ	n	n	p	p	b	b	f	f	v	v	m	m	.
VOCALIC	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
CONSONANTAL	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
COMPACT	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
GRAVE	-	-	-	-	-	-	-	-	-	-	-	+	+	+	+	+	+	+	+	+	+	+
NASAL	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	-	-	+	+	+
SHARP	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+	-	+
CONTINUANT	-	-	-	-	+	+	+	+	-			-	-	-	-	+	+	+	+			
VOICED	-	-	+	+	-	-	+	+	-			-	-	+	+	-	-	+	+			
STRIDENT	-	-	-	-					+													

TABLE B. THE PHONEMES OF RUSSIAN
re-ordered to eliminate the ambiguous zero

by the number of zeros eliminated from Table A. Although it is different for different phonemes, on the average it is 6.5 bits per phoneme, a value considerably lower than our original 11 and nearer to the ideal value of 5.38. The description in terms of features has thus been made efficient.

4. CONSIDERATIONS OF PHONEME PROBABILITIES

(a) Individual Frequencies of Occurrence

The next step in our description of the language will be to consider the relative frequencies of the individual phonemes. The 'information' in bits per phoneme obtained previously has the hypothetical minimum value of 5.38 ($\log_2 42$), a

result obtained by successively subdividing the set of phonemes into two equal groups. When their frequencies of occurrence are unequal, however, the required average (bits per phoneme) is obtained by successively subdividing the set into two groups of equal total probability. The result then is that the average number of questions needed to identify a phoneme (in terms of bits per phoneme) is H_1 , where

$$H_1 = -\sum p_i \log p_i \quad (1)$$

summed over all phonemes i . This is the 'expected' value of $-\log p_i$. (Remember that p_i is always less than 1).

<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>e</i>
a	1316	2.94	.387	4	d	177	5.81	.100	9
i	977	3.35	.328	6	l,	162	5.95	.096	4
t	602	4.05	.244	9	'u	153	5.96	.091	6
'a	539	4.23	.228	4	r,	133	6.20	.083	4
j	457	4.45	.202	2	z	130	6.25	.081	8
n	392	4.66	.183	6	d,	126	6.30	.080	9
'o	379	4.72	.179	5	b	119	6.39	.075	8
s	359	4.80	.172	8	x	102	6.60	.067	5
'e	343	4.86	.167	5	g	91	6.80	.062	7
k	284	5.14	.146	7	v,	89	6.84	.061	8
v	273	5.15	.140	8	3	89	6.84	.061	6
'i	243	5.38	.131	6	f	85	6.86	.058	8
u	240	5.40	.129	6	s,	85	6.86	.058	8
p	232	5.42	.126	8	š	59	7.40	.044	9
r	230	5.45	.125	4	m,	56	7.50	.043	6
n,	221	5.50	.121	6	b,	52	7.60	.039	8
l	212	5.55	.118	4	p,	50	7.64	.038	8
ʃ	207	5.56	.115	6	k,	36	8.10	.029	7
m	202	5.64	.114	6	z,	21	8.90	.018	8
c	197	5.65	.111	5	f,	8	10.30	.008	8
t,	196	5.65	.111	9	g,	7	10.50	.008	7

TABLE C

a = Phoneme (i); $b = p_i \times 10^4$; $c = -\log_2 p_i$; $d = -p_i \log_2 p_i$;
 e = number of features listed in Table B (i means 'any given phoneme'; p_i means 'the probability of a given phoneme')

The relative frequencies of the individual Russian phonemes have been counted from samples of the language, as described in Section 1; they are listed in Table C. From these frequencies p_i we may readily calculate the hypothetical 'information' H_1 given by (1). This is

$$H_1 = 4.78 \text{ bits/phoneme} \quad (2)$$

On the other hand we may calculate the average number of features, i.e. the binary choices per phoneme, knowing the probabilities p_i of the individual phonemes. If N_i is the number of features required to identify the i th phoneme in Table B, then the rate of feature choices which actually occurs is

$$\sum N_i p_i = 5.79 \text{ features/phoneme} \quad (3)$$

which may be compared to the ideal given by (2).

In a recent article,¹⁰ Huffman has described a method for devising the most efficient code possible for a set of independent messages of known frequency distribution. In such a code 'the average number of coding digits per message is minimized'. If we regard the phonemes of our language as independent messages, we can apply Huffman's method and compute from the probabilities given in Table C the number of digits which in an optimal code would be necessary to identify each phoneme uniquely. This can be compared to the number of features necessary to identify each phoneme in Table B. It must be pointed out, however, that these are not strictly comparable: as we stated in Section 3, the description in terms of distinctive features presupposes that the digits are interpreted differently depending on the answers given in a preceding stage of the analysis, while in Huffman's code all digits have the same interpretation. In the following table we compare the number of phonemes having a given number of digits in the optimal code with the number of phonemes having the same number of distinctive features in Table B.

NUMBER OF DIGITS OR DISTINCTIVE FEATURES										
	2	3	4	5	6	7	8	9	10	
In an optimal code	0	2	2	11	13	8	3	1	2	
In the actual case (Table B)	1	0	6	4	10	4	12	5	0	

Regarded purely as a descriptive process, then, the method of listing the distinctive features is rather efficient.

So far we have been regarding the phonemes of the language as independent. But the natural process of speech consists not merely of choosing a chain of independent phonemes; at the very least it consists of a succession of choices, where each choice is in part conditioned by the preceding phoneme chosen. It may be a truer description of the natural process of speech to say that phonemes are chosen in groups. Thus the simple analysis that we have made so far must be regarded as a somewhat artificial though quite efficient description of the language in its simplest aspect.

Before concluding this section on individual phonemes, it may be of interest to note a few statistical facts gathered from Table C.

Probability of a vowel occurring = 0.4190; of a liquid = .0737; of a glide /j/ = 0.0457; of a consonant proper = 0.4616.¹¹

In the accompanying table, the plus and minus probabilities of each feature were calculated by adding the probabilities of all phonemes showing a plus for that feature in Table B and of all those showing a minus. Thus the probability

¹⁰ David A. Huffman, A method for the construction of minimum redundancy codes, *Proceedings of the IRE* 40:9.1098-101 (1952).

¹¹ Markov, in his study of LETTER distributions in a Russian poem, obtained the value 0.4317 for vowels and 0.5683 for consonants. His figures are remarkably close to ours, especially if we make allowance for the fact that Markov counted some instances of /j/ as vowels and others not at all.

of a yes-answer to the question Voiced? is the sum of the probabilities of /g-g,-z-d-d,-z-z,-b-b,-v-v,/, while the probability of a no-answer is the sum of the probabilities of /k-k,-t-t,-s-s,-p-p,-f-f,/. (We omit the data concerning nasality, stridency, compactness, and diffuseness; for here the pluses are much fewer than the minuses, and the lower probability of the former is obvious.)

	PROBABILITY OF	
	+	-
Voiced	.1174	.1920
Sharp	.1242	.3445
Stressed (vowels only)	.0935	.2533
Continuant	.1822	.2530
Grave (vowels)	.0772	.1563
Grave (consonants)	.1684	.2861
Totals	.2456	.4424

These figures are significant, especially since the pluses and minuses were assigned without considering their relative frequency, entirely on the basis of an examination of the features and their interrelations.¹²

But the phonemic structure of a language is not defined entirely by the total probabilities of feature occurrence; their distribution in time is also significant. These distributions measure what might be termed the continuity of each feature; they can be obtained from the analysis of joint probabilities presented below. Thus, if we know the probabilities $p(a\ b\ c\ \dots\ n)$ of various chains of n phonemes, we can readily assess the probability that a certain distinctive feature exists uninterrupted for a duration greater than m phonemes, where $m = 1, 2, \dots n$. It is not our purpose here to execute such an analysis in detail, but rather to point out its potentialities as a basis for language description.

(b) *Phoneme Groups, Syllables*

In the preceding section we paid attention mainly to what may be termed phonemic monograms—that is, to individual phonemes, with some reference also to phoneme groups and to their joint probabilities of occurrence. These groups may be digrams, trigrams, and so on. Another type of probability which is of interest to the student of language structure is the TRANSITION PROBABILITY that a particular phoneme will follow a given phoneme or phoneme group. Thus, if $p(a\ b\ \dots\ n)$ is the probability of the phoneme group $(a\ b\ \dots\ n)$, then

$$\begin{aligned}
 p(a\ b\ \dots\ n) &= p(a)p_a(bc\ \dots\ n) \\
 &= p(a)p_a(b)p_{ab}(cd\ \dots\ n) \\
 &= p(a)p_a(b)p_{ab}(c)p_{abc}(d\ \dots\ n), \text{ etc.}
 \end{aligned}
 \tag{4}$$

In this way the joint probability of a group is related to the transition probabilities of the successive phonemes a, b, c , etc. occurring in the group.

Given a particular phoneme (a) of a language, or a possible group of phonemes

¹² For a fuller discussion, see R. Jakobson, *Sound and meaning* (to appear).

(ab ... n), the phonemes (m) which can occur next in the chain have a set of probabilities $p_{ab...n}(m)$. The fact that these probabilities vary according to the character of m implies that a certain degree of prediction is possible. This property provides another form of 'redundancy' in the language, a quality which is of great importance in aural recognition, as when we follow a conversation in a noisy room.

For instance, if one hears a palatalized /v, / in a Russian utterance, one can be sure that no unstressed vowel except /i/ will follow. After a palatalized /b, /, the probability of an unstressed /a/ is extremely low; the sequence /b,a/, as in /g^lolub,a/ 'pigeon' (gen.-acc. sing.) and /gal^lub,a/ 'fondling', is exceptional. In our count we have found the following phonemes after palatalized /s, /, with the indicated frequencies:

i	33	^l a	6	u	1
^l e	16	m	2	j	1
^l o	9	^l u	1	a	1
^l i	8				

Note especially the almost complete absence of consonants and the very low frequency of unaccented /a/. On the other hand, after nonpalatalized /s/ the unaccented /a/ was the most frequent of all the vowels in our material, and consonants occurred very freely. Our figures for phonemes after /s/ are these:

t	76	p	9	k,	3
a	37	u	6	j	2
t,	30	v	6	^l i	1
k	27	i	5	m,	1
l	20	m	5	r	1
^l a	16	n,	5	^l u	1
^l o	11	p	5	v,	1
l,	10	x	3	r,	1
n	10				

Since the inequality of the transition probabilities makes possible a certain degree of prediction, the information conveyed by one phoneme in the chain of connected speech is less than that conveyed by one phoneme in isolation. Unless it is the first in the chain, we know something about it, so to speak, before it arrives. This information can be strictly defined, in the technical sense of the earlier sections; we can even derive formulae, analogous to equation (1), which will be applicable to connected groups of phonemes. Suppose, for example, that we have computed the probabilities $p(ab)$ of all the phoneme digrams of a language; then the information conveyed by any digram of the language is, on the average, $H_{1,2}$:

$$H_{1,2} = -\sum p(ab) \log p(ab) \text{ bits/digram} \quad (5)$$

Similarly for trigrams:

$$H_{1,2,3} = -\sum p(abc) \log p(abc) \text{ bits/trigram} \quad (6)$$

But if, instead, we have computed the various transition probabilities $p_a(b)$, the information conveyed by the occurrence of each successive phoneme is $H_1(2)$:

$$H_1(2) = -\sum p(ab) \log p_a(b) \quad (7)$$

Again, if we know the transition probabilities $p_{ab}(c)$:

$$H_{1,2}(3) = -\sum p(abc) \log p_{a,b}(c) \quad (8)$$

Clearly these various information rates, based on different probability tables, are connected. To show this, consider equation (4); take logs of both sides and then average over all possible groups ($ab \dots n$):

$$\begin{aligned} & -\sum p(ab \dots n) \log p(ab \dots n) = \\ & -\sum p(ab \dots n) [\log p(a) + \log p_a(b) + \log p_{ab}(c) \dots] \text{ or} \\ H_n &= H_1 + H_1(2) + H_{1,2}(3) + H_{1,2,3}(4) \dots \text{ bits/n-gram} \end{aligned} \quad (9)$$

This means that the information conveyed by groups of phonemes is, on the average, equal to the sum of the information obtained from each successive phoneme.

We have computed the values for the digrams and trigrams in our material according to the first count—the one that takes account of the boundaries between words and between the parts of compounds. The values were found to be 8.45 bits/digram and 9.15 bits/trigram. If the phonemes were independent, the corresponding values would be 9.54 bits/digram and 14.31 bits/trigram. As expected, the values are lower when the units in the chain are not regarded as independent.

Another very promising approach, which for the present must remain unexplored, is to calculate the distributions of the distinctive features in time, as already proposed in Section 4(a). Given a long sample of text transcribed phonemically, we write under each symbol a column of pluses, minuses, and zeros representing its distinctive features in some regular order (as in Table A). The horizontal sequences of pluses, minuses, and zeros produced in this way can then be used to measure the 'continuity' of the various features. The probabilities of such sequences may be written $p_+(m)$, $p_-(m)$, $p_0(m)$, where $m = 1, 2, 3$, etc. It is obvious that such distributions may provide a basis for statistical specification of the phonemic differences between one language and another.

The statistical analysis of the phonemes and their sequences in connected messages must be supplemented by a similar analysis of the dictionary, in order to understand the distribution of phonemes in the lexical code of the given language.¹³ The comparison of the two sets of data is certain to be most instruc-

¹³ In R. Carnap's terminology, the occurrences of phonemes, having been studied in the Russian word-EVENTS, are to be investigated in the word-DESIGNS, just as we have here studied the occurrences of distinctive features in the phoneme-DESIGNS; cf. *Introduction to semantics* 3 (Cambridge, Mass., 1946). Charles S. Peirce, the founder of modern semiotic, would say that besides the application of the phonemic LEGISIGNS within the lexical SINISIGNS, such an application must be scrutinized again within lexical LEGISIGNS; cf. his *Collected papers* 2.245-7 (Cambridge, Mass., 1932).

tive. The statistical analysis of the dictionary permits us to draw conclusions about the phoneme sequences peculiar to different types of morphemes and to words of different grammatical categories.¹⁴ Furthermore, it forms the basis for definitive statements about phoneme combinations with probabilities of 1 and 0; for no phoneme sequence can occur in messages if it is not provided by the code.

Finally, among problems which remain to be investigated are those transitional probabilities which operate backwards, i.e. which depend not on earlier but on subsequent events, or, in linguistic terms, not on progressive but on regressive action of phonemes in a sequence. The comparison of these two sets of statistics is very important, because it is obvious that for different types of sequence the predictability is greater in one direction than in the other. Analysis of such data will provide the most solid basis for setting up a statistical model of the syllable as a recurrent link in the chain of speech.

¹⁴ An exhaustive statistical analysis of the phonemic structure of Russian root morphemes is being prepared by Robert Abernathy within the framework of the research program mentioned in fn. 2.

A QUASI-ARITHMETICAL NOTATION FOR SYNTACTIC DESCRIPTION

YEHOSHUA BAR-HILLEL

Massachusetts Institute of Technology

The purpose of this paper* is to present the outlines of a method of syntactic description that is new insofar as it combines methods developed by the Polish logician Kasimir Ajdukiewicz¹ on the one hand and by American structural linguists² on the other. Such a combination has apparently not been undertaken before, if only for the reason that Ajdukiewicz's paper appeared in a Polish philosophical journal and has therefore remained unknown to most linguists.

We are not interested here in developing a method which a linguist might use to ARRIVE at an analysis of a linguistic corpus, but only in a new way in which he could PRESENT the results of his investigations. The decisive difference between this method and the others prevailing so far lies in the fact that in addition to a list in which each linguistic element (usually each word) is assigned to one or more categories, only a simple rule of a quasi-arithmetical character need be given to enable us to 'compute' the syntactic character of any given linguistic string (sequence of one or more elements) in its context. The main economy produced by this method lies, therefore, in that it enables us to dispense completely, at least in principle, with special syntactic statements.

This should be of value in those situations in which a completely mechanical procedure is required for discovering the syntactic structure of a given string. Such a situation arises, for instance, in connection with the problem of mechanized translation. It has been shown³ that a machine could be constructed which would be able to determine the structure of any sentence in the source language, provided that the syntax of this language were presented to the machine in a certain specific form which we may call OPERATIONAL. It may well be that the preparation of the element-category list will be a decisive step toward the construction of an operational syntax whose instructions could be carried out satisfactorily by a digital computer or, for that matter, by a human being operating in a completely mechanical fashion.

Before we proceed to sketch the new method in abstracto, we will discuss one simple example. The English string *Poor John sleeps* would be analyzed, accord-

* This work has been supported in part by the Air Materiel Command, the Signal Corps and the Office of Naval Research, and in part by a grant from the Rockefeller Foundation.

¹ Die syntaktische Konnexität, *Studia philosophica* 1.1-27 (1935). An English mimeographed translation, under the title Syntactic Connection, appeared at the College of the University of Chicago, March 1951.

² As presented, for instance, by Zellig S. Harris, *Methods in structural linguistics* (Chicago, 1951), or by Charles C. Fries, *The structure of English* (New York, 1952).

³ Yehoshua Bar-Hillel, The present state of research on mechanical translation, to appear in *American documentation*.

ing to one method⁴ recently described, in the following way: *poor* is an A (for adjective), *John* is an N (for noun), *sleep* is a V (for verb), *-s* is a Vv (for morpheme added to a verb to form a verbal phrase), where all these assignments hold at least for the given context. Then, by invoking the syntactic statements, $A \frown N \rightarrow N$ (where the arch designates concatenation and the arrow stands for yields) and $V \frown Vv \rightarrow V$, *Poor John* would be assigned to N, *sleeps* to V, hence finally *Poor John sleeps* to $N \frown V$ which is the most frequent form of English sentences.

According to the notation to be proposed and explained in this paper, *John* will belong to the category n , *poor* to $\frac{n}{[n]}$, *sleeps* to $\frac{s}{(n)}$, where n is to be interpreted, approximately, as the category of name-like strings, $\frac{n}{[n]}$ as the category of those strings that with an n to their right form a string belonging to the same category n , and $\frac{s}{(n)}$ as the category of those strings that with an n to their left form a string belonging to the category of sentences. That the string *Poor John sleeps* is a sentence can now be tested mechanically, without recourse to any syntactic statements, by using something like ordinary arithmetical multiplication of fractions on the INDEX-SEQUENCE corresponding to the given string, viz.

$$(1) \quad \frac{n}{[n]} n \frac{s}{(n)}.$$

By REDUCING the sub-sequence $\frac{n}{[n]} n$ to n , we obtain the FIRST DERIVATIVE

$$(2) \quad n \frac{s}{(n)},$$

from which, by another reduction, we get the SECOND and LAST DERIVATIVE

$$(3) \quad s.$$

Let us notice immediately another important advantage of our notation: we have not only a mechanical method of testing the SYNTACTIC CONNEXITY of a given string but also a mechanical method of finding the so-called constituents of any given syntactically connex string. In the given example, we find by quick inspection that *John sleeps* is not a constituent of *Poor John sleeps* (in spite of the fact that *John sleeps* is connex in itself), since reducing first $n \frac{s}{(n)}$ to s we would arrive at the derivative

$$(2') \quad \frac{n}{[n]} s,$$

from which no further derivation is possible, showing (by the fact that the last derivative, or EXPONENT, in this case does not consist of a single index) that the whole string, analyzed in this way, is not syntactically connex. Strictly speaking, we are only entitled to the following conditional statement: IF *Poor John sleeps* is a syntactically connex string, then *John sleeps* is not one of its constituents.

⁴ See *Methods*, ch. 16. Harris uses an equal-sign instead of the arrow, and juxtaposition instead of the arch.

Let us also take notice that a complete element-list of the stated character enables us to synthesize all possible sentences of the given language without any additional rules. Let us consider, for instance, a language that contains elements belonging, respectively, to the categories n , $\frac{n}{[n]}$, $\frac{n}{[s]}$, and $\frac{s}{(n)[n]}$ (the category of strings that form sentences out of a left n and a right n). Then we know that any syntactically connex sequence of elements belonging to n , $\frac{s}{(n)[n]}$, $\frac{n}{[s]}$, n , $\frac{s}{(n)[n]}$, $\frac{n}{[n]}$, $\frac{n}{[n]}$, n , in this order, would be a sentence, since the only possible exponent of the corresponding index-sequence is s , as can be seen from the following derivation (using, for the sake of typographical simplicity, a slant-line instead of the fraction-bar):

- | | | | | | | | | |
|-----|-----|------------|---------|-----|------------|---------|---------|-----|
| (4) | n | $s/(n)[n]$ | $n/[s]$ | n | $s/(n)[n]$ | $n/[n]$ | $n/[n]$ | n |
| (5) | n | $s/(n)[n]$ | $n/[s]$ | n | $s/(n)[n]$ | $n/[n]$ | | n |
| (6) | n | $s/(n)[n]$ | $n/[s]$ | n | $s/(n)[n]$ | | | n |
| (7) | n | $s/(n)[n]$ | $n/[s]$ | | | s | | |
| (8) | n | $s/(n)[n]$ | | n | | | | |
| (9) | | | | | | | s | |

Whether all these categories have instances in a given language is of course an empirical question. With respect to English, for instance, this question could be answered Yes, roughly speaking. Indeed, *John knew that Paul was a poor man* would be a sentence of the type mentioned. (The qualification 'roughly speaking' is necessary, since it is obvious that the categories dealt with so far are too gross to be applicable to ordinary languages; according to such an application, *Man knew that John was poor a Paul* would also have to be considered a sentence, a thing which most people would hesitate to do.) The same holds for French, where we have sentences like *Jean savait que Paul était un pauvre homme*, but not for German, which lacks elements belonging to the category $\frac{n}{[s]}$ (at least on an unsophisticated level, for which *Paul war ein armer Mann* and *Paul ein armer Mann war* are not automatic alternates of the same sentence, but which considers the first string alone as a sentence and the second as syntactically disconnex).

These preliminary considerations should suffice to show that the designation of syntactic categories with the help of such symbols (which carry with them, so to speak, indications of the environments in which members of these categories appear) is a method certainly superior to other prevailing ones in at least one respect, if it can be carried through consistently. Rudiments of this notation already appear in the symbol V_v that was used in our first illustration to designate the category to which $-s$ belongs.

Let us begin by stating some assumptions on which our method is based. We assume that with respect to any two given ELEMENT-TOKENS (ink-marks, sounds) it is known whether or not they stand in the relation of EQUIFORMITY, which we take in this context as an undefined primitive relation: Whenever two element-tokens are equiform we shall say that they belong to the same ELEMENT-FORM. The relation of equiformity can be extended, in an obvious way, to hold also between STRING-TOKENS. If all tokens of some form belong to the same

category, then this form will be called *PURE*. If not all tokens of some form belong to the same category, but each token belongs to exactly one of n categories, we shall call the form *MIXED* and consider it as the (set-theoretical) sum of n *TYPES* belonging to n different type-categories. A pure form is, of course, its only type. To illustrate: assuming that all tokens of *Paul* belong (in English) to the same token-category (viz. of proper names), then the type *Paul* belongs to the type-category of proper names. Assuming that some tokens of *poor* belong to a certain category, others to another category, still others to a third one, and none to another, we shall consider the form *poor* as mixed and composed of the three types, say *poor*₁, *poor*₂, and *poor*₃.

Observing a given token, we know, in general, to which form it belongs, but not, unless the form is pure, to which type. To find out the type of a token in a given context, we usually have to take account of the linguistic environment of the token and often also of the extra-linguistic context of its production. In all those cases where linguistic environment alone is sufficient to fix the type of a given token, our notation will facilitate this determination and formalize it in such a way that a suitably constructed machine should be able to carry it out. (Even this statement needs qualification: it seems that the proposed notation will be effective only where the crucial environment is not too extended. When the elements are words, the environment taken into account in our notation does not go beyond an utterance.)

According to the envisaged notation, to each element-form there will be assigned a class of n (≥ 1) symbols denoting the categories of the types to which the tokens of this form belong. These symbols will be called the *INDICES* of the form and their class the *index-class* of this form. To a sequence of elements the sequence of their index-classes will be correlated. To arrive at the category-system, we make, among others, the following assumptions. Each sentence that is not an element is regarded as the outcome of the operation of one sub-sequence upon the remainder, which may be to its immediate right or to its immediate left or on both sides. ('Left' and 'right' are to be understood here, as in what follows, only as the two directions of a linear order.) That sub-sequence which is regarded as operating upon the others will be called an *OPERATOR*, the others its *ARGUMENTS*. In a two-element sentence, for instance, one element will have to be the operator, the other its argument. In this case, having only one argument, the operator is *SINGULARY*. In other cases the operator may be *BINARY*, *TERNARY*, etc. According to the position of the arguments, we have to distinguish between a singular right operator, singular left operator, binary right operator, binary left operator, binary right-left operator, etc. The verbal terminology will soon become pretty involved. We shall therefore use in general the following symbolism: an operator that forms a sentence out of m left arguments belonging (from right to left) to the categories $\alpha_1, \dots, \alpha_m$, respectively, (the α 's may be different but need not be so) and out of n right arguments belonging (from left to right) to the categories β_1, \dots, β_n , respectively, will be said to belong to the category

$$\overbrace{(\alpha_m) \dots (\alpha_1)[\beta_1] \dots [\beta_n]}^s \quad (m + n \geq 1)$$

If both the operator and the arguments are elements, then the only remaining thing to be done is to assign the arguments to such categories as will yield an over-all simplest description. (The tremendous problems connected with this procedure cannot be discussed here.) If, however, either the operator or some argument, or both, are PROPER STRINGS, i.e. consist of more than one element, they too have to be regarded as the result of the operation of some sub-sequence upon the remainder. The result of the operation in this case will, in general, no longer be a sentence. Whenever an operator, out of m left arguments belonging to $\alpha_1, \dots, \alpha_m$ and n right arguments belonging to β_1, \dots, β_n , respectively, forms a string belonging to the category γ (identical with one of the α 's or β 's or different from all of them), it will be said to belong to the category

$$\frac{\gamma}{(\alpha_m) \dots (\alpha_1)[\beta_1] \dots [\beta_n]} \quad (m + n \geq 1)$$

Elements which are operators in a given string may be arguments in another string or even at another place in the same string, and similarly with arguments. It seems plausible, however, that the requirement of over-all greatest simplicity will necessitate, at least with respect to languages with a finite number of elements, the classing of some types as arguments in all contexts. We shall now make the assumption that the languages we are dealing with contain types which are 'absolute' arguments, so to speak. These types will be said to belong to BASIC CATEGORIES. For purposes of illustration, we shall assume that sentences, proper names, and common nouns belong to basic categories. As the symbol for the category of sentences we shall continue to use s , if necessary with subscripts. As the symbol for the other basic categories we shall use n —again, if necessary, with subscripts. (If the language to which these categories apply is not assumed to be fixed in the given discourse, superscripts can be used to indicate that language relative to which the categorization is to hold.)

Strings which belong to basic categories will themselves be called BASIC. Strings that are not basic are operators, and belong to OPERATOR-CATEGORIES. There is a potentially infinite and elaborately ramified hierarchy of them, but in an effective description of a given language only some of them will be used.

We are now ready to describe that operation upon index-sequences, to be called DERIVATION, on the basis of which the other concepts of our method will be defined. By a derivation of an index-sequence we understand the replacement of any sub-sequence of the given sequence of the form

$$\alpha_m \dots \alpha_1 \frac{\gamma}{(\alpha_m) \dots (\alpha_1)[\beta_1] \dots [\beta_n]} \beta_1 \dots \beta_n \quad (m + n \geq 1)$$

by γ . The resulting index-sequence is called the DERIVATIVE of the original sequence.

A given index-sequence may obviously have more than one derivative. We already saw that the index-sequence (1) had both (2) and (2') as its derivatives. Another example is provided by the string *a very poor man*, to which, among

others, the following index-sequence is correlated (the use of the double slant-line should be sufficiently clear):

$$(10) \quad n/[n] \quad n/[n]/[n/[n]] \quad n/[n] \quad n$$

This has two derivatives:

$$(11) \quad n/[n] \quad n/[n] \quad n$$

$$(11') \quad n/[n] \quad n/[n]/[n/[n]] \quad n.$$

A derivative may, in its turn, have one or more derivatives. Thus, (11) has the following as its only derivative:

$$(12) \quad n/[n] \quad n,$$

which again has as its only derivative

$$(13) \quad n,$$

whereas (11') has no derivative at all. (11) and (11') may be called the **FIRST DERIVATIVES** of (10), (12) its **SECOND DERIVATIVE**, (13) its **THIRD DERIVATIVE**, (11') and (13) its **LAST DERIVATIVES OR EXPONENTS**. (13), but not (11'), is a **PROPER EXPONENT**, i.e. one consisting of a single index, and any derivation leading up to it a **PROPER DERIVATION**. The terms 'derivative' and 'exponent' will be used not only with respect to index-sequences but also with respect to the strings to which these sequences are correlated.

Since an element may have more than one index correlated to it, a string may have more than one index-sequence correlated to it. If at least one index sequence of the set of index-sequences correlated to a given string has at least one proper exponent, the string will be called (**SYNTACTICALLY**) **CONNEX**. *Poor John sleeps* and *a very poor man* are both connex, each having at least one proper derivation; *poor sleeps John* is not, since neither of its two index-sequences

$$n/[n] \quad s/(n) \quad n$$

$$n \quad s/(n) \quad n$$

has a proper derivation, as can be verified immediately.

Here, however, arises the following interesting situation. A string that is connex by itself need not be **CONNEX AS A SUB-SEQUENCE OF SOME OTHER STRING**. Before we define this phrase, let us give an example. *John sleeps* is connex by itself but is not connex within *Poor John sleeps*, so far on an intuitive basis.

Let us now give the strict definitions: A string m_1 will be said to be **CONNEX AT A CERTAIN PLACE WITHIN A STRING m_2 WITH RESPECT TO THE DERIVATION d_1** if (1) m_2 is connex, (2) d_1 is proper, (3) d_1 includes a subderivation in which the index-sequence of m_1 at the place in question has a proper exponent. An exact definition of the term 'subderivation' would be somewhat tiresome; it is hoped that the illustration given will make its intended meaning sufficiently clear. In the proper derivation (1)—(2)—(3) above, the index-sub-sequence correlated to *Poor John*, viz. $n/[n] \quad n$, has the exponent n . *Poor John* is therefore connex

within *Poor John sleeps* with respect to this derivation, but *John sleeps* is not connex within *Poor John sleeps* with respect to this derivation.

We now define ' m_1 is connex within m_2 ' as short for ' m_1 is connex within m_2 with respect to all proper derivations of m_2 ', and ' m_1 is thoroughly connex' as short for ' m_1 is connex within all m_2 of which it is a (proper or improper) part'.

Clearly, not every connex string has to be also thoroughly connex. In English, *John sleeps* is connex but not thoroughly connex since it is not connex within *Poor John sleeps*. That a language should exhibit this character may be deplored, since it introduces complications into its description and into the analyses carried out on the basis of such a description. We shall take up this point again at a later stage.

The complications mentioned are not such as to cause, by necessity, any major ambiguities. The knowledge that a string is thoroughly connex would indeed dispense with the task of testing whether this string is connex within some given context. That this knowledge is not at our disposal might necessitate more complex checking procedures, but the outcome of these procedures can still be unique. Knowing that *John sleeps*, though connex, is not thoroughly connex, we might be interested in finding out whether it is connex within *Paul thinks that John sleeps*, or at least whether it is connex within this larger string with respect to some of its proper derivations. This last question can indeed be answered in the affirmative by exhibiting the following proper derivation:

	Paul	thinks	that	John	sleeps
(14)	n	s/(n)[n]	n/[s]	n	s/(n)
(15)	n	s/(n)[n]	n/[s]	s	
(16)	n	s/(n)[n]	n		
(17)		s.			

The relevant subderivation is framed. But is *John sleeps* also connex with respect to all other proper derivations of *Paul thinks that John sleeps*? The derivation given above is the only proper one with (14) as the original index-sequence. But (14) is only one out of many other possible original sequences. *Thinks* may also have at least the indexes s/(n) and s/(n)[s] (as in *Paul thinks* and *Paul thinks John is sleeping*, waiving possible sophistications) and *that* also has the indexes n and n/[n] (as in *Paul believes that* and *Paul likes that girl*). Disregarding other possible indexes, we have therefore before us at least nine original index-sequences for the given string, which we might arrange in the following way:

Paul	thinks	that	John	sleeps
	s/(n)	n		
n	s/(n)[s]	n/[s]	n	s/(n)
	s/(n)[n]	n/[n]		

By systematic testing we can find that only one other original index-sequence

out of the possible nine has a proper derivation. The sequence and its derivation are:

(14') n s/(n)[s] n/[n] n s/(n)

(15') n s/(n)[s] n s/(n)

(16') n s/(n)[s] s

(17') s.

Since *John sleeps* is not connex within *Paul thinks that John sleeps* with respect to this derivation, it is not connex within this larger string (without qualification). In this case, however, we can describe the situation in a slightly more precise way and say that *John sleeps* is connex within *Paul thinks that John sleeps* with respect to a certain index-sequence of this larger string, since *John sleeps* is connex within *Paul thinks that John sleeps* with respect to every proper derivation starting with this original sequence.

The fact that (14') is a far less likely sequence for *Paul thinks that John sleeps* than (14) is of high importance for what we might call STATISTICAL SYNTAX. Our investigation lies on that level where relative frequencies are not yet taken into account, but only possibilities and impossibilities of occurrence. We also disregard, on this level of approximation, the differences in intonation-patterns which would determine, with high likelihood, whether a given uttered token of *Paul thinks that John sleeps* has the one or the other of the two stated original sequences assigned to it.

Instead of the phrase 'is connex within m_2 with respect to d_1 ' we shall, in general, use the more customary expression 'is a constituent of m_2 with respect to d_1 ', hence instead of 'is connex within m_2 ' also 'is a constituent of m_2 '. We can now also define the phrase 'is an immediate constituent of m_2 with respect to d_1 ' as meaning 'is a constituent of m_2 with respect to d_1 but is not a constituent of any m_3 that is a (proper) constituent of m_2 with respect to d_1 '. In our last example, for instance, *that John sleeps* is an immediate constituent of *Paul thinks that John sleeps* with respect to the derivation (14)—(17), and also with respect to the derivation (14')—(17'), and this in spite of the fact that it has different exponents in these two derivations.

With respect to the first derivation, the immediate constituents of *Paul thinks that John sleeps* are *Paul*, *thinks*, and *that John sleeps*. Of these, the first two are elements and the third a proper string which has, therefore, immediate constituents of its own, with respect to the same derivation, viz. *that* and *John sleeps*, of which the first is an element and the second again a proper string with the immediate constituents *John* and *sleeps*. In a self-explanatory terminology, we may therefore say that, with respect to the given derivation (14)—(17), the elements *Paul* and *thinks* are IMMEDIATE CONSTITUENTS of the given string, *that* is a CONSTITUENT OF THE SECOND ORDER, *John* and *sleeps* are CONSTITUENTS OF THE THIRD ORDER; and we may say that the string itself is OF THE THIRD ORDER. A connex two-element string would then be of the first order with respect to any

of its proper derivations, and we might say, if this should prove to be convenient, that any element is OF ZERO ORDER (without qualification).

If two tokens of the same element-form occur in a given string, then explicit reference to these occurrences may have to be made, since even if they belong to the same type they need not be constituents of the same order with respect to a given proper derivation. It is clear that m_1 may be an immediate constituent of m_2 with respect to some proper d_i but not with respect to some different d_j . If, however, m_1 happens to be an immediate constituent of m_2 with respect to all proper derivations, then we shall drop the qualifications and say that m_1 is an immediate constituent of m_2 . Under the assumption that the set of index-sequences given above for *Paul thinks that John sleeps* is exhaustive, *Paul*, *thinks*, and *that John sleeps* are immediate constituents of this string, without qualifications.

Many of the concepts defined so far are much more dependent than one might at first thought assume upon the specific derivation, and hence upon the form of the original index-sequence. One might tend to believe, for instance, that it would make no appreciable difference whether one regarded a given operator as forming a sentence out of a right n and a left n , i.e. as an operator $s/(n)[n]$, or as forming out of a right n an operator that forms a sentence out of a left n , i.e. an operator $s/(n)//[n]$. But such a belief would be a mistake. It makes a considerable difference in the organization of immediate and other constituents whether we treat *loves* (say) as an operator which out of a left n *John* and a right n *Mary* forms a sentence, *John loves Mary*, IN ONE COMPLEX STEP, or as an operator which out of a right n *Mary* forms an operator, *loves Mary*, which out of a left n *John* forms a sentence IN TWO SIMPLE STEPS. According to the second treatment, *loves Mary* is an immediate constituent of the whole; according to the first, it is no constituent at all. The fact that being-a-constituent-of is a relation which is not invariant even with respect to such 'inessential' transformations as that of $s/(n)[n]$ into $s/(n)//[n]$ shows that this relation and its cognates are of somewhat restricted importance. Incidentally, according to the Aristotelian analysis, *John loves Mary* has to be understood as a subject-predicate sentence with *John* and *loves Mary* as its immediate constituents. The categorization which treats *John* and *Mary* on a par, as respectively the left and right arguments of *loves*, though much in favor with modern logicians, for a long time was not regarded as proper.

But leaving philosophical and logical considerations aside, it is an interesting problem to compare the advantages and disadvantages of using only SINGULARY OPERATORS WITH COMPLEX NUMERATORS as against using only n -ARY OPERATORS WITH SIMPLE DENOMINATORS, or using both simultaneously. In certain presentations of COMBINATORIAL LOGIC,⁵ for instance, singulary operators are preferred, in spite of the fact that this involves multiplying the number of operations. Nothing more will be said here on this topic.

It is useful, in certain investigations, to distinguish between operators which

⁵ Cf. Haskell B. Curry. A theory of formal deducibility, *Notre Dame mathematical lectures* No. 6, 1950.

out of their arguments form a string belonging to the same category as the arguments, and those which do not. The first kind might be called **ENDOTYPIC**, the second **EXOTYPIC**. That type of *poor*, for instance, which belongs to the category $n/[n]$, is endotypic, while *sleeps*, which belongs to $s/(n)$, is exotypic. In certain contexts, it might be profitable to use a slightly different classification and to regard operators belonging to categories of the form $\alpha/(\alpha) \dots (\alpha)[\alpha] \dots [\alpha]$ as endotypic. That type of *and*, for instance, which is an $s/(s)[s]$, would be endotypic according to the second conception but exotypic according to the first. We might perhaps, if necessary, distinguish between endotypic in the narrower sense and endotypic in the wider sense.

English adjectives when used in adjectival function, demonstratives when used in adjectival function, articles, adverbs, and conjunctions are endotypic; verbs and prepositions are exotypic; nouns in general are neither, being mostly basic. This is only a rough application of our terminology, since it is obvious that it is unable, as developed so far, to cope with the whole gamut of relationships that exist between the elements in English or in any other natural language. With this provision in mind, adjectives will in general belong to n/n -categories (omitting parentheses and brackets now for the sake of simplicity), as will articles and adjectival demonstratives; conjunctions will be s/ss (hence endotypic only in the wider sense); adverbs will be $n/n/n/n$ (*VERY good*), $s/n/s/n$ (*sleeps SOUNDLY*), $s/nn/s/nn$ (*ARDENTLY hates*), s/s (*UNFORTUNATELY John died*), etc.

To get a slightly better outlook on the effectiveness of the proposed notation, let us analyze a string with a much more complex structure than that of the strings so far, though still very far from the top of the complexity ladder. The string is taken from a language⁶ which most readers will not know. Its rough transliteration is:

1	2	3	4	5	6	7	8	9
moše	yada	ki	pinzas	xaxam	yoter	meašer	axoto	haktana

Let us assume that by looking up the elements of these strings in the category-list (so far non-existent), we obtain the information given in Table 1. (The list in Table 1 is incomplete even with respect to that very rough approximation which we have set as our standard.) We have before us a set of 216 original index-sequences. Systematic testing for **SUITABLE** original sequences (i.e. sequences with proper derivations) would be laborious though perfectly feasible for a properly designed machine. We shall use some shortcuts. It is easy to realize that 9a does not fit. 5a does not fit either. Similarly 5b, 6c, hence also 7c, are out. The exponent of 7-9 obviously operates on the exponent of 4-6 and must be *s*, because of 3. 3b is unsuitable, as is 2a. Exactly one suitable original index-sequence is left:

	1	2	3	4	5	6	7	8	9
(20)	n	$s/(n)[n]$	$n/[s]$	n	$s/(n)$	$\frac{s/(n)}{(s/(n))}$	$\frac{s/(n)}{(s/(n))[n]}$	n	$n/(n)$

⁶ Colloquial Hebrew. For the sake of the simplicity of analysis, however, the common *meqzoto* has been replaced by the (colloquially) much less frequent *meašer axoto*.

Many derivations start from (20). There are already three different first derivatives as a result of operating 5 upon 4, 6 upon 5, and 9 upon 8. It is, however, easy to see that the first derivation leads into a blind alley. We thus arrive at the interesting (though on second thought not surprising) result that only two proper derivations are correlated to the given string, even though it has 216 original index-sequences, many of them with more than one first derivative. We present one of the two proper derivations in the abbreviated scheme of Table 2. The second derivation differs from the one presented in Table 2 only in that steps

	1	2	3	4	5	6	7	8	9
a	n	$s/(n)$	$n/[s]$	n	n	$\frac{n/(n)}{(n/(n))}$	$\frac{n/(n)}{(n/(n))[n]}$	n	n
b		$s/(n)[n]$	$s/[s]$		$n/(n)$	$\frac{s/(n)}{(s/(n))}$	$\frac{s/(n)}{(s/(n))[n]}$		$n/(n)$
c					$s/(n)$	$\frac{s/(n)[n]}{(s/(n)[n])}$	$\frac{s/(n)[n]}{(s/(n)[n])[n]}$		

TABLE 1

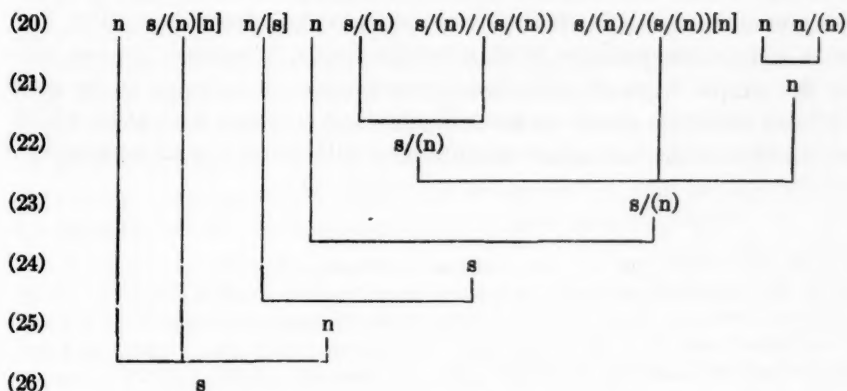


TABLE 2

(21) and (22) change places, a difference which we may safely describe as trivial. If the category-list were constructed so as to contain singular indexes only, the derivation would have contained two more steps, each THREE-FORK being split into a sequence of two TWO-FORKS, and the set-up of immediate and n -order constituents would have undergone some changes.

I am fully aware of the inadequacies of the proposed notation. Its shortcomings are many and of various kinds. Some are due to the effort of presenting the main ideas as simply as possible, and can easily be overcome through a less simplified approach. Moreover, the problem of categorization is far from a satisfactory solution, and the proposed notation as such (I repeat) cannot put the linguist in a better position for solving it, though it may well redirect his attitudes.

Two further points should be mentioned. First, we have said nothing about what the linguistic elements are to be in specific cases; in our examples, we used words. It is plausible that other elements might be more suitable, if not for all languages, at least for many. Phenomena like separable prefixes and Harris's 'long components' will pose additional problems.

Another feature, related to the previous, is our deliberate restriction of the range of arguments to the IMMEDIATE environment of the operators. This will prove to be disturbing with regard to languages that have constructions like the English string *Paul strangely enough refused to talk*, where *strangely enough* is probably best regarded as an s/s operator, since it is the whole event which is regarded as strange and not the way in which Paul refused to talk. But whether we assign s/(s) or s/[s] to this string, it will turn out that the larger string will be disconnex, as the reader can verify for himself. However, both *Paul refused to talk* (,) *strangely enough* and *Strangely enough* (,) *Paul refused to talk* will come out all right. The last two sentences are commonly regarded as variants of the first, with hardly any difference in meaning—certainly not in cognitive meaning.

Though this example points to a certain shortcoming of our notation, it also shows that only a small change is needed to make it adequate for handling complications of this type. Using certain auxiliary symbols such as commas would change the picture. If we write *Paul, strangely enough, refused to talk* (which is, incidentally, the common usage), and interpret the function of the commas as giving us license to lift the string between them from its position and deposit it at some other position (within certain limits, of course), we can still adhere to the simple rules of immediate environment. It remains to be seen whether devices of such a simple nature will enable us to retain a notation which takes account only of the immediate environment with respect to all languages.

MISCELLANEA

CONTRAST

BERNARD BLOCH, YALE UNIVERSITY

Contrast between sounds can be defined, I think, on the basis of distribution alone, without the customary appeal to meaning. On two earlier occasions I tried to formulate such a definition,¹ but both times without success—largely because the essential criterion was not yet clear to me. A number of readers, fastening on the weak spot in my argument, have quite properly refused to accept the conclusions based on it.² This third attempt to state my position has the excuse that I now think I know what the position is.

The range of an allophone³ in some finite sample of a dialect (theoretically also in the dialect as a whole) is defined by a list of all the environments in which members of the allophone occur. Less precisely but more usefully, the range of an allophone can also be defined in general terms, by reference to the features of position, accent, or surrounding sounds that are common to the environments of the member phones or to particular subsets of member phones, but without specific mention of any individual environments. What such a GENERAL DEFINITION defines is not the exact range of the allophone in question, but rather a class of possible environments, including the range of the allophone and also some environments in which members of the allophone do not occur. By making the number of features referred to in the definition large enough, it is possible to reduce the otiose environments to a practicable minimum, a point at which their existence does not impair the diagnostic value of the definition for the purposes of phonemic analysis. But although these otiose environments can be reduced in variety, they cannot be eliminated altogether so long as the definition of the range does not list the constituent environments one by one.⁴

¹ A set of postulates for phonemic analysis, Lg. 24.3-46 (1948), esp. 22-6, Postulates 25-31; and the definitions prefixed to an account of Japanese phonemes, Lg. 26.89-90 (1950).

² See Eli Fischer-Jørgensen, Remarques sur les principes de l'analyse phonémique, TCLC 5.213-34 (1949); John Lotz, Speech and language, Journal of the Acoustical Society of America 22.712-7 (1950); Kenneth L. Pike, More on grammatical prerequisites, Word 8.106-21 (1952).

³ For the definition of technical terms, see the two passages cited in fn. 1. Specific references: for *allophone*, Lg. 24.38 (§54.2); for *dialect*, 24.8 (§3.2); for *environment*, 24.22 (§25.3), 26.89; for *phone*, 26.89, and cf. 24.12 (§12.3), 24.35 (§52.3); for *range*, 24.30 (§42.3).

⁴ Thus, the range of [ñ] (the class of all voiced nasalized apical flaps or one-tap trills) in some sample of English can be very simply defined: members of the allophone occur after a vowel or semivowel and before a weak-stressed vowel (as in *winter*, *painting*, *counter*, *carpenter*, *Mount Olympus*). The range of [ŋ'] (the class of all front or prevelar voiced nasals), again in some sample of English, requires a general definition somewhat more complex: members of this allophone occur (1) final in a syllable after a front vowel (as in *bingo*, *sang*, *making*) or after the semivowel [j] (as in *boing* or a monosyllabic pronunciation of *being*, *saying*, *trying*); (2) after a front vowel or [j] before prevelar [k'] in a strong-stressed syllable (as in *think*, *winked*, *jinx*, *Schenck*, *Jenks*, *bank*, *Manx*, *Schumann-Heinck*, *oink-oink*); and (3) after the vowel [ɛ] before [b] in a strong-stressed syllable (as in *length*). If the choice of defining features is correct, any environment in which a member of one of the

Comparing the range of one allophone (however defined) with the range of another allophone in the same dialect or sample, yields a statement of the relation between the two ranges, or the **RELATIVE DISTRIBUTION** of the two allophones. For any pair of allophones P and Q, there are four possible types of relative distribution:

(1) **COMPLEMENTARY**: the ranges of P and Q are mutually exclusive. If the environments are listed in full, the two lists have no member in common; if the ranges are defined in general terms, the two definitions have no feature (of position, accent, or surrounding sounds) in common.

(2) **COINCIDENT**: the ranges of P and Q are the same. Every environment of P is also an environment of Q,⁵ and conversely.

(3) **INCORPORATING**: the whole range of P is a part of the range of Q. That is, the range of P is wholly included within the range of Q, but not conversely. Every environment of P is also an environment of Q, but Q has environments in which P does not occur.

(4) **OVERLAPPING**: a part but not all of the range of P is a part but not all of the range of Q. The two ranges have some environments in common, but each range includes environments which are not in the other.

If we speak of the environments common to a pair of ranges as the (*pars*) **COMMUNIS**, and of the environments in one range not shared by the other as the (*pars*) **PROPRIA**,⁶ we can characterize the four types of relative distribution as follows: in complementary distribution, there is no *communis*; in coincident distribution, there is no *propria* in either range; in incorporating distribution, one range has a *propria* but the other has not; in overlapping distribution, each range has a *propria* in addition to the *communis*.

The rest of this paper will be stated in postulational form.⁷

Postulate. In any dialect there are pairs of allophones P and Q in overlapping or incorporating distribution, such that the *communis* of their ranges cannot be distinguished from either *propria* by any general definition.

Corollary. For any such pair of allophones P and Q, the *communis* of their ranges can be distinguished from either *propria* only by a complete listing of the individual environments.

Definition. Such allophones P and Q are *in contrast* with each other, and any member of P is *in contrast* with any member of Q.

allophones occurs will be found to exhibit one of the features referred to in the definition. But there will be countless environments, both among those that actually occur in the sample studied and among the vastly greater number that do not occur but can be invented, which exhibit one of these features and yet contain no member of the allophone in question.

⁵ Provided that each environment is interpreted as composed of allophones rather than phones. The statement in the text can be more pedantically paraphrased as follows: To any environment of P consisting of the phones $a_1b_1c_1 \dots$ there corresponds an environment of Q consisting of the phones $a_2b_2c_2 \dots$, such that a_1 and a_2 , b_1 and b_2 , c_1 and c_2 , etc., belong respectively to the same allophones—and conversely.

⁶ These terms are adopted from Rulon S. Wells, *Automatic alternation*, Lg. 25.99–116 (1949), esp. 104. (Wells uses the terms *communis* and *propria* for parts of morphs.)

⁷ This postulate is intended to replace the seven postulates referred to in fn. 1. Other parts of my 1948 set of postulates are similarly in need of restatement and simplification.

Definition. Any pair of allophones P' and Q' to which the provisions of this postulate do not apply are *noncontrastive* with each other, and any member of P' is *noncontrastive* with any member of Q'.

LONG AND SHORT IN ICELANDIC PHONEMICS

KEMP MALONE, JOHNS HOPKINS UNIVERSITY

This paper is restricted to one feature of the Icelandic sound system: quantity in the structure of the syllable. Moreover, I am confining myself to strest syllables; quantity in unstrest syllables will not be taken up. My treatment of the material will be synchronic, current Icelandic alone being considered.

The strest syllables of modern Icelandic fall into four types, as they do in English. In type A the syllable consists of a sonant (i.e. syllabic phoneme) only; no consonants (i.e. asyllabic phonemes) occur. In the other three types, the syllable is made up of a sonant and one or more consonants. In type B the sonant is preceded, in type C followed, in type D both preceded and followed, by one or more consonants. Examples:

type A: *t* 'in'type B: *bj* 'bee'; *frt* 'holiday'type C: *ts* 'ice'; *trsk* (adj.) 'Irish'type D: *dts* 'goddess'; *djpt* 'depth'

Coming now to the matter of quantity in these syllabic types, we can dispose of some points at once. A sonant that makes (type A) or ends (type B) its syllable is always long. A single consonant that begins its syllable (types B and D) is always short, as is also any member of an initial consonant group. A sonant that does not end its syllable (types C and D) may be either long or short; so also a single consonant that ends its syllable. Here, however, only two sequences occur: short sonant plus long consonant, and long sonant plus short consonant. A consonant group that ends its syllable serves as the equivalent of a single long consonant, and the preceding sonant is therefore short. Examples of short sonant plus long consonant: *kerti* 'candle', *menn* 'men', *heimta* 'claim'; examples of long sonant plus short consonant: *ker* 'vessel', *men* 'necklace', *heim* 'home'; examples of short sonant plus consonant group: *heimska* 'folly', *danska* 'Danish', *fyrst* 'first'.

Long consonants that occur between sonants make a special case; examples: *amma* 'grandmother', *verra* 'worse', *gadda* 'spike'. In such words the bulk of the consonant goes with the preceding sonant, giving the familiar syllabic pattern: short sonant plus long consonant. The very end of the consonantal articulation, however, goes with the following sonant and the consonant is thus split in two, part belonging to one syllable, part to another. Such consonants are often called double; the conventional spelling answers to an important phonetic feature of their articulation.

How are we to interpret these pronunciations phonemically? Some have argued that the short sonants are allophones of the same phonemes as the long, since they occur only before a long consonant or before a consonant group,

positions in which the long sonants are not found. There can be no doubt that short and long sonants are in complementary distribution. But the short and long consonants are likewise in complementary distribution: a consonant is always short unless it ends a syllable; a consonant that ends a syllable is long if it comes immediately after a short sonant; otherwise such a consonant is short. The principle of complementary distribution here fails us. It seems clear that in such an opposition as *men* 'necklace' : *menn* 'men', the long *e* is opposed to the short and the short *n* is opposed to the long. It will not do to play favorites here, settling upon the opposition *n* : *nn* as phonemic and reducing the opposition *e* : *ē* to allophonic status. There is no logical basis for choosing one opposition rather than the other. It is therefore not only simpler but also sounder to say that the distinction between long and short is phonemic for both sonants and consonants.

I can see but one other possibility. We may abstract quantity from the individual phonemes and take it to be a feature of the syllable, a feature comparable to stress and pitch. If we do this, identifying length in a phoneme with quantitative accent, we see at once that an Icelandic strest syllable takes such an accent on its surgent.¹ In syllabic types A and B this coincides with the sonant. In types C and D surgent and sonant may or may not be identical. If the surgent is a consonant it takes the quantitative accent in spite of its subordination in terms of sonority. In other words, the surgent of an Icelandic strest syllable is regularly long; the other phonemes of the syllable are regularly short. We may therefore use the conventional sign of length to mark stress as well. In such an opposition as *mēn* 'bracelet' : *meñ* 'men', the macron tells us not only which phoneme is long but also which phoneme is the surgent.

¹ For the term *surgent* see my *Phonology of modern Icelandic* 5, 6, and passim (1923). Here it will be enough to define the surgent of a syllable as the phoneme that coincides with the intensity peak.

ADDENDA TO THE COMPARISON OF INEQUALITY IN SPANISH

DWIGHT L. BOLINGER, UNIVERSITY OF SOUTHERN CALIFORNIA

This extension of remarks on the subject of my article, *The comparison of inequality in Spanish*, Lg. 26.28-62 (1950), has two chief motives. First is my inexcusable oversight in failing to collate Hayward Keniston's article, *Expressions for than after a comparative in sixteenth century Spanish prose*,¹ with my own data. While Keniston treats another epoch, it is important to note that the two descriptions are nearly identical, save for two sharp deviations noted below. Second is the discovery of an apparent contradiction which turns out to be predictable on the basis of the theory and hence confirms it.

1. The first of the two respects in which contemporary Spanish differs radically from 16th-century Spanish was noted in the original study (31 and note 21). It is the reversal of *no más de*, the earlier stereotype for 'not other than' = 'only', reserved now for the true quantitative 'not more than' and replaced in the sense of 'only' with *no más que*. Keniston's article (148), however, gives the impression

¹ *Revue de linguistique romane* 6.129-51 (1930).

of a sharper cleavage than was noted there on the basis of his *Syntax of Castilian prose* §26.345 and §26.541 (Chicago, 1937), leading one to posit a reorientation of *que* and *de*: the *que* of 'otherness' has been generalized, and the *de* of 'fromness', which Keniston regarded as the natural method of indicating 'no different from', hence 'only', has shrunk to the orbit of straight quantification, analogous to English *upwards of*. The two are exemplified in the following citation from Palacio Valdés:²

Hace media hora que ustedes están juntos, y las reglas de la casa no permiten más que quince minutos ... Ninguno puede estar junto a una niña más de ese tiempo.

The second is that contemporary Spanish with very few exceptions avoids the type *Habla más que sabe* 'He does more talking than knowing', frequent in 16th-century Spanish,³ in favor of the type *Habla más de lo que sabe* 'He talks more than what he knows' (cf. my article, note 55).

2. The apparent contradiction is evidenced in the following (Grismer-Arjona 29):

Yo os perdono, y más aún, hija mía, te felicito por haber escogido un marido mucho mejor del que yo hubiera podido darte.

According to the formula, this should have *que el que*, but my two Spanish colleagues prefer *de*. There is clearly no question of confusing *mayor* with *más*, as I noted (my article 48) with previous instances of interchanged *que* and *de*.⁴ The native speaker's first guess is euphony, but this is obviously the usual failure to see 'sounding good' as an effect rather than a cause. The test for the example cited lies in making the second term refer to a definite individual: when I changed the subordinate clause to *el que yo le había escogido*, both speakers shifted their preference to *que*. In the following they reacted according to the standard pattern, demanding *que* in all:

Este aparato es más complicado que el que tengo en mi laboratorio — Estas señoras son mayores que las que viven en el otro piso — Estas reliquias son más antiguas que las que nos mostraron en Roma

But in the last example, when the second term was changed to *las que hubiéramos podido hallar en Roma*, they accepted *de*. In the following, where the second term is vague, *de* was acceptable; but as soon as a definite, concrete individual or thing replaced the hypothetical one, they again demanded *que*:

Este reducto es más fuerte del que habría sido posible construir en la colina [but que el que se construyó, etc.] — Sus inventos son más ingeniosos de los que otro hubiera podido idear — La vida del reconcentrado ha sido más terrible de la que cualquiera haya soñado en las pesadillas del delirio — Una fortuna más grata de la que nunca nos imagináramos

Thus when the second member is vague, the speaker ceases to think of counter-

² Raymond L. Grismer and Doris King Arjona, *The pageant of Spain* 122 (New York, 1938).

³ Keniston, *Expressions for than* 138.

⁴ Cf. this additional example of *mayor* for *más*: *mayor mal del que me ha sucedido no puede sucederme* (G. A. Becquer, *Legends, tales, and poems* 40 [ed. Olmstead; New York, 1907]); and the invented example *La parte es menos que el todo*, judged possible by a Castilian speaker alongside of *menos de* (but *menos de*, not *que*, *mi porción*, *parte*, *cuota*, etc.).

balancing two different things, and reverts to the more powerful suggestion of a different degree of the same thing. If it is true that *que* yields to *de* under the impulse of indefiniteness, then *de* should yield to *que* under the impulse of excessive definiteness. One condition of definiteness is singularity. In *Tiene más autoridad de/que la que tiene un príncipe*, if *autoridad* is construed as a mass noun the choice is *de*; if as a countable, *que*: four native speakers—a Spaniard, a Costa Rican, an Argentinian, and a Peruvian—preferred *que* here but accepted *de*. The extreme case is where we have a purely quantitative comparison, with the second term a countable in the singular naming a precise object. The utter quantitateness demands *de*, the utter definiteness demands *que*, and apparently no regular construction with an explicit second member is possible. For example, if only one man attended a dance on a previous occasion, we cannot use the Spanish equivalent for *There are more men at the dance than (supply he who) attended last time*. With the slightly less precise but still quantitative ‘one vote’, *que* would be used:

Hubo más votos por el candidato republicano que el solo voto [*but de los 1.347 votos*] *que* recibió en las elecciones anteriores.

A reexamination of the examples shows that the polarity definite-indefinite also helps the confusion in the formal likeness of *menor-menos*, *mayor-más*: note the use of *que* with the singular ‘one’ in *mayor que uno*, against *mayor de treinta* (48).⁵

Another condition of definite and precise ‘otherness’ in which the pattern is overbalanced is found in the speaker’s intent to exaggerate. In *Las manos eran, en efecto, tan bellas, más bellas que lo que D. Luis había dicho en sus cartas*,⁶ Valera enhances the contrast more than with the lukewarm *de lo que*. In *Este* [sic-miners’ usage] *agua sube más que lo que me dijo* (invented by L. Moreno), the speaker might be expressing angry surprise, or calling the other person a liar. In *El matrimonio pesa más que lo que yo creía* (similarly), the speaker implies that marriage is something quite different from what he had expected. The effect achieved is predictable from the regular pattern. *Que* normally substantivizes a following *lo que*, thus contrasting two names of separate things rather than two aspects of the same thing.

In the following example,⁷ a singular is used in the same vague, undifferentiated way as the plurals above:

Me parece que no se me puede tener en peor opinión de la que ya he logrado. The ‘otherness’ of two distinct opinions is submerged in the gradation of a single opinion; the expression is equivalent to *no puede ser peor de lo que es*.

3. ‘As has happened with English *different*, certain other expressions have by their close similarity in meaning been made similar to comparisons in form as well’ (29).⁸ The Academy (§428j) notes this for *diferente*, *distinto*, *diverso*, and *primero*. It extends, however, all the way to the comparison of identity, as might

⁵ Bello (§1021) implies that *mayor (menor) de veinticinco años* is elliptical for *que de*. The contamination described here is probably a factor.

⁶ Pepita Jiménez 124 (Clásicos castellanos).

⁷ J. Figueroa Campos, *El misterio de los guantes negros* 30 (Buenos Aires, 1947). This reference, and others marked GPS, were supplied by G. P. Sullivan.

⁸ Subsequent references are to pages of my earlier study.

be expected, since *mismo* likewise takes the same correlative *que*. Here it is with the antonym of *mismo*, *contrario*:⁹

Con las mujeres ocurre lo contrario que con las lanchas de salvamento: nos gustan más las de los otros que las que consideramos como nuestras
From there it passes to fractional and multiple comparisons, e.g. *Tiene doble (triple, la tercera parte, etc.) que antes*, or, with *mitad*:¹⁰

El tipo de cambio actualmente es muy ventajoso y los libros les resultan a mitad de precio que antes

The multiples analogize even farther: *Tiene doble de lo que debe tener ~ Tiene más de lo que debe tener*. But with *veces* the comparison is expressed as an out-right multiplication: *Tiene dos veces lo que debe tener*.

The Academy's argument from ellipsis (§428j), that *acaban de comer con costumbres diferentes que equals diferentes de las costumbres con que*, is weakened by the type *El suyo es muy diferente que el mío*, accepted by my Castilian colleagues, where no ellipsis is possible. There is likewise no relevant ellipsis in *amar a una persona de distinta patria que la nuestra*.¹¹ As it is more usual in English to employ *than* where *from* would be wordy (e.g. *It caused quite different reactions in a man than [from those which it caused] in a woman*¹²), so Pereda writes:¹³

La madre iba por caminos diferentes [de los caminos por los] que su marido.
The vacillation between *que* and the prepositions *de* and *a* is seen in this instance of *a* with a true comparative and in the following *otro de*:¹⁴

Las palabras españolas terminadas en consonante representan un promedio de un 38%, en tanto que en inglés ese mismo tipo de vocablos suele aparecer ... en proporción no menor al 67% — Es, evidentemente, la vida más otra de la nuestra que cabe imaginar

4. 'As with English *prefer*, *preferir* may govern *que*' (29). Numerous written examples now confirm the oral one cited, for instance these:¹⁵

Prefiero aburrirme con la cartera llena que con la carta vacía — Góngora prefiere ser condenado por ignorante que por hereje

So for *ser preferible*:¹⁶

Es preferible gustar las cosas hora por hora ... que pasar vertiginosamente por la vida

5. The assimilation of *antes* and *después* to the comparatives in *antes que* and *después que* (rather than *antes de que* and *después de que*) was noted in my earlier paper (30). This is corroborated by the use with them of the quantifier *mucho*, which, like English *much*,¹⁷ is employed with regular comparatives (*mucho mayor*, *much older*). An example:¹⁸

⁹ Enrique Jardiel Poncela, *Amor se escribe sin hache* (Buenos Aires, 1945). GPS.

¹⁰ Letter from León Sánchez Cuesta, 7 July 1949.

¹¹ Antonio Azorín, *Old Spain* 82 (New York, 1928). GPS.

¹² Advertisement in *Collier's* for 12 August 1939, p. 48.

¹³ J. M. Pereda, *Peñas arriba* 1.82 (Buenos Aires, 1942). GPS.

¹⁴ Tomás Navarro, *Fonología española* 49 (Syracuse, 1946); and J. Ortega y Gasset, *Revista de occidente* 5.30 (1927). The latter reference supplied by Joseph Silverman.

¹⁵ Azorín 32; and Agapito Rey, *Hispania* 32.403 (1949).

¹⁶ Azorín 56.

¹⁷ *Much* is also a test for the assimilation of *prefer* to the comparative: *I much prefer it*, but not *I much like it*.

¹⁸ Samuel Gili Gaya, *Revista de filología española* 30.116 (1946).

comienzan a evolucionar las formas compuestas mucho antes que en castellano

6. Besides the type *La gruta me confunde más de lo [confuso] que estoy* (49), where *más* modifies a combination of verb and adjective instead of the more customary single element, one may get other overlappings like *Lo hace con más frecuencia de lo que creo conveniente*, where if the single element were modified one would expect *de la que*. Obviously *con frecuencia* functions as a unit equivalent to *frecuentemente*, and the sentence is construed *más con frecuencia de lo que*, etc.¹⁹ But in *Hay más distancia de lo que yo creo conveniente*²⁰ the noun is apparently lost sight of.²¹

7. *Suficiente* occurs in a transposed order, *Tenemos más que suficientes razones para rechazarlo*, which seems to parallel *perdió más de diez dólares* (45) and hence to call for *de*. Native speakers prefer *que*, however, and the following examples²² show that the phrase was inverted, and equivalent to *Tenemos razones más que suficientes*, not to *Tenemos más razones que suficientes*:

un torrente circulatorio imperfecto, un hígado tórpido ... son motivos más que suficientes para que el cerebro de un anciano no trabaje con la regularidad debida — Fué este un motivo más que suficiente para que incurriera en el odio profundo de los abolicionistas

For *de* to be acceptable, *suficiente* has to be made quantitative rather than descriptive by adding the article: *Tenemos más de las suficientes razones*.

¹⁹ The parallel example from Keniston, *Syntax of Castilian prose* §26.441, is *lo tal creo con más veras de lo que se me puede decir*, where *más* 'logically modifies the adverbial phrase *con veras*'.

²⁰ Pío Baroja, *Paradox*, *Rey* 51 (New York, 1937).

²¹ Keniston, *Syntax of Castilian prose* §26.441, cites *porque no uuiesse dentro del Alhambra más mal de lo que podía aver con la gente que avía dentro*.

²² *Revista de los archivos nacionales de Costa Rica* 12:1/2.52; *ibid.* 48.

AN ELAMITE ETYMOLOGY

HERBERT H. PAPER, CORNELL UNIVERSITY

This note¹ discusses an Elamite etymology, hitherto unrecognized, to show how certain hypotheses concerning the interpretation of Elamite cuneiform writing can lead to a deeper understanding of Elamite phonology and morphology, and to a more precise identification of particular forms known from Elamite texts. Among the Elamite documents, the royal texts from the Achaemenid period represent a particular dialect that we shall call ROYAL ACHAEMENID ELAMITE (RAE).² The existence of a large number of graphic variants for the same form in the RAE texts led Weissbach, as early as 1890, to certain phonemic conclusions regarding the Elamite cuneiform symbols.³ He recognized the non-

¹ The substance of this paper was presented at the Cincinnati meeting of the American Oriental Society in April 1950. The etymology proposed here has been accepted by the following Elamite scholars: George Cameron, Johannes Friedrich, Richard Hallock, and René Labat.

² For a summary and a rough dialect division see Labat, *Structure de la langue élamite*, *Conférences de l'Institut de linguistique de l'Université de Paris* 9.23-42 (1951).

³ F. H. Weissbach, *Die Achämenideninschriften zweiter Art* 31 (Leipzig, 1890).

phonemic character of written double consonants, and indicated the correct interpretation of the broken-vowel writing ($CV_1V_2C = CV_1(V_2)C$).

With these two basic phonological assumptions in mind, we shall examine the connection between the RAE forms *an-la-gi* and *ap-pan-la-ik-ki-um-me*.

Bisitun paragraph 18 contains the following phrase:⁴ RAE *ša-u-mi-in* ^d*u-ra-maš-da-na* ^h*ti-ik-ra an-la-gi ú-ud-da* = OP *vašnā Auramazdāha Tigram viyatarayāmā* = Bab. *ina GIŠ.MI (= šilli) šá* ^d*Ú-ri-mi-iz-da* ^{Id}*Di-iq-lat ni-te-bi-ir*. The Old Persian and Babylonian versions mean, 'By the grace of Ahuramazda, we crossed the Tigris'. What is expressed in these versions by *viyatarayāmā* and *nitebir* 'we crossed' is rendered in RAE with the noun-and-verb phrase *an-la-gi ú-ud-da*, lit. 'I made a crossing'. The form *an-la-gi* is a hapax, but its meaning is supplied for this context by the other versions; *ú-ud-da* is an orthographic variant of the better known *hu-ud-da*, first person singular declarative of the verb-base /ʔut-/ 'to make'.

The form *ap-pan-la-ik-ki-um-me* occurs three times in RAE (Bisitun §63 bis, §64). According to the hypotheses mentioned above, this form can be interpreted as /apanlakime/, the double consonants being equivalent to single, *la-ik* to *lak*, and *ki-um* to *kim*. Its contexts are the following:⁵

Bisitun §63: RAE *a-ak in-ni* ^v*ib-ba-ak-ra in-ni* ^v*iš-tuk-ra ap-pan-la-ik-ki-um-me hu-ud-da* = OP *naiy šakauði[m naiy ...]nuwatam zūra akunavam* = Bab. *a-na liq-tu ù muš-ki-nu [par]-ki ul e-pu-[uš]*. That is, 'Neither to an X nor to a Y did I do a wrong-doing'. And again in the same paragraph: RAE *ap-pan-la-ik-ki-um-me ak-ka-ri-ug-gi in-ni hu-ud-da*, with no counterpart in the OP and Babylonian versions; it may be translated, 'I did not do a wrong-doing to anyone'.

Bisitun §64: RAE ^v*nu* ^v*sunki* ^v*ak-ka me-iš-ši-in ša-ni-ik-ti* ^v*ruh* ^h*ir-ra ti-tuk-ra hu-pir-ri a-nu in kan-ni-in-ti a-ak a-nu* ^v*ak-ka ap-pan-la-ik-ki-um-me hu-ut-ti-iš* 'You who will later become king, a man a liar, him do not befriend nor one who does wrong-doing'. The Babylonian has the fragmentary *par-ka-ni la ta....*, whereas the OP has no equivalent for this passage. From its OP counterpart *zūra*, it seems clear that *ap-pan-la-ik-ki-um-me* means 'wrong, wrong-doing'.

I suggest that *an-la-gi* and *ap-pan-la-ik-ki-um-me* are closely related. The latter is analyzed as consisting of the so-called inanimate relative pronoun /apa/, /anlaki/ 'a crossing, a going-over' (quotable in the form *an-la-gi*), and /me/, the abstractive suffix which turns *sunki* 'king' into *sunki-me* 'kingship'. Thus, *ap-pan-la-ik-ki-um-me* may be literally translated 'what-beyond-ness' and therefore 'wrongdoing, transgression'. The semantic relationship of *an-la-gi* to *ap-pan-la-ik-ki-um-me* is precisely the equivalent to Akkadian *nabalkuttum*, Hebrew *abērah*, Greek *παπαβαίνω*, Latin *trānsgridior*.

The recognition that *an-la-gi* and the kernel of *ap-pan-la-ik-ki-um-me* are one and the same form has the value of showing still another graphic variant to exemplify the two conventions of Elamite graphemics already alluded to: the

⁴ Weissbach, *Die Keilinschriften der Achämeniden* 24-5 (Leipzig, 1911). Weissbach's transliteration is here modified in the following respects: superscript *d*, *h*, *v*, *lg* represent respectively the Sumerian DINGIR-sign, the single horizontal cuneiform stroke, the single vertical cuneiform stroke, and the Sumerian MEŠ-sign, which merely marks the logographic use of the preceding sign.

⁵ Ibid. 66-7.

non-phonemic nature of written double consonants and of the second vowel in broken-vowel writings. Still a third point is involved; namely, that the consonantal value of the GI-sign is the same as that of the KI-sign.

It is clear that a systematic descriptive study of Elamite is a prime desideratum and must precede any attempt to investigate the possible relationship of Elamite to other languages.⁶ Until a good deal more is known about this language, all discussions of possible relationship are fruitless.

⁶ Cf. my dissertation (unpublished), *The phonology and morphology of Royal Achaemenid Elamite* (University of Chicago, 1951).

REVIEWS

The mathematical theory of communication. By CLAUDE L. SHANNON and WARREN WEAVER. Pp. vii, 117. Urbana: University of Illinois Press, 1949.

Reviewed by CHARLES F. HOCKETT, *Cornell University*

Most of this book (1-91) consists of an article by Shannon, bearing the same title as the volume, which first appeared in the Bell System Technical Journal for July and October 1948. The remaining section, by Weaver, is entitled Recent contributions to the mathematical theory of communication; a more condensed version appeared in the Scientific American for July 1949. Weaver's paper is less technical than Shannon's, and might well have been placed first in the book, as an introduction to Shannon's exposition. In this review explicit references to Weaver will be few, but this is deceptive: the reviewer found Weaver's more discursive treatment of great value in grasping Shannon's often highly technical presentation, and the reader who chooses to pursue the subject further will do well to read Weaver's paper before attempting Shannon's.

A number of other contributions to the theory of communication have appeared in recent years. Two by Robert M. Fano are worth mentioning here: The transmission of information, Technical reports No. 65 (17 March 1949) and No. 149 (6 February 1950) of the Research Laboratory of Electronics, Massachusetts Institute of Technology. Fano's discussion is helpful because of a difference of approach, though his results are substantially the same as Shannon's.¹

The appearance of the term 'communication' or 'information' in the title of an article or a book is naturally no guarantee that its contents are of any concern to linguists. Shannon's work stems in the first instance from engineering considerations—telegraph, teletype, telephone, radio, television, radar, and the like—which would seem rather remote. But the theory is rendered so general in the course of mathematicizing that it may turn out to apply, in part or with

¹ These and other contributions to information theory refer constantly to the work of Norbert Wiener: the famous *Cybernetics* (1948), and *The extrapolation, interpolation, and smoothing of stationary time series with engineering applications* (1949; earlier printed as an NDRC Report, MIT, 1942). *Cybernetics* consists of chapters of extremely difficult prose alternating with chapters of even more difficult mathematics; the other volume is reported to consist almost completely of the latter. The reviewer had managed to absorb some odd bits of the prose parts of the first of these before Shannon's articles appeared; the mathematical parts are entirely beyond his capacity.

A fairly popular discussion of some aspects of information theory will be found in E. C. Berkeley, *Giant brains* (1949), particularly the earlier chapters.

In June and July 1951, a grant from the Committee on the Language Program of the ACLS enabled the reviewer to attend the first intensive summer course at MIT on Modern Communications; the various lectures and discussions in this course were of considerable help in following the work of Shannon and others in this field. However, the reader should be warned that the reviewer's training in mathematics is very slight, and that as a result parts of this review may be based on serious misunderstandings.

some modifications, to communication via language; this, at least, is a possibility that must be investigated.²

We divide our review into three sections. In the first we outline what seem to be the key points in Shannon's theory. In the second we discuss linguistic applications. In the third we take up certain more general issues.

THE THEORY

'The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point. Frequently messages have *meaning*', but this is irrelevant for the communications engineer. 'The significant aspect is that the actual message is one *selected from a set* of possible messages. The system must be designed to operate for each possible selection, not just the one which will actually be chosen since this is unknown at the time of design' (3). Shannon's aim is to find some method for quantifying the commodity that is carried by a communications system, in order to establish a valid measure of the efficiency of such systems, and, primarily, to establish theoretical limits to this efficiency in terms of the known or determinable variables. The practical use of such theoretical limits is to channel design-research towards attainable aims. The commodity itself Shannon calls INFORMATION.

The keynote of the quantification of information is the matter of choice of any message, for actual transmission at a given time, from a fixed repertory of possible messages. If you are constrained to answer every question I may put to you with a simple 'yes', then there is no point in my asking any questions at all; I know in advance what the answer will be. If a telegraph key is tied down so that the receiver buzzes constantly, ENERGY is being expended which would not be expended if the key were open for the same indefinitely long period, but no more INFORMATION can pass from transmitter to receiver under the first condition than under the second. Somewhat more subtly, if the value of the mathematical constant π is computed at one geographical point, and the result, to any desired number of decimals, is transmitted to another geographical point, no information has been sent: the value of π is determinate, so that all one would have to do at the second point would be to perform the computations independently; the answer would be the same (31). Thus, if there is no indeterminacy, no element of choice, there can be no transmission of information.

If your answer to my questions, on the other hand, may be either a simple 'yes' or a simple 'no', with no further alternatives—as in the game Twenty Questions—then the system will transmit information from you to me, still providing that I do not know in advance which answer you will give to each question. If, in addition to 'yes' and 'no', there is some third alternative, say 'maybe', it seems logical to say that the system can transmit more information

² References to information theory have already appeared in linguistic discussions: J. Whatmough, presidential address before the Linguistic Society, December 1951; C. F. Hockett, Lg. 27.337, 445 (1951); R. Jakobson, C. G. M. Fant, and M. Halle, *Preliminaries to speech analysis*, Technical report No. 13 (January and May 1952) of the Acoustics Laboratory, MIT. It is not certain that all these references are based on adequate understanding of the theory.

per message, on the average, than with only two alternatives. That is, the larger the repertory of possible messages, the larger, in general, is the informational CAPACITY of the system.

For various reasons the measure actually chosen is not the raw count of the number of messages in the repertory, but rather the logarithm of this number to the base 2—providing that the number of messages in the repertory is finite, and that they are all equally likely to be chosen. If either or both of these conditions is lacking, then the measure of amount of information becomes more intricate; but the more complicated formula reduces to the one described above when the conditions are met.³

The unit of information thus measured and quantified is called the BINARY DIGIT, BIGIT, BINIT, or BIT; the last term is currently the most favored, but we shall use BINIT.⁴ A term is needed for the unit of capacity; we define one SHANNON as a capacity of one binit of information per second.

Thus, in the scheme outlined above, where I ask you questions which must be answered with 'yes' or 'no', and where those answers are equally probable, we have a system with a capacity of one binit per message. The fundamental assumption in the game Twenty Questions is that any animal, vegetable, or mineral can be specified unambiguously, on the average, by twenty successive dichotomizings of the appropriate kingdom; that is, that twenty binit of information will usually suffice for such specification. Skill at interrogation in this game consists in so phrasing one's questions that the region specified by the answers to all previous questions is divided into two essentially equiprobable subregions.

However, there is an important but peculiar restriction on the use of information-theoretical methods. They serve to measure the entropy of an information-source or the capacity of a channel (the terms will be defined in a moment), but they afford no means whereby we can state how much information is actually conveyed by the actual selection and transmission of any specific message. Your equiprobable yesses and noes transmit ON THE AVERAGE one binit of information each but how much information is carried by any one specific yes or no is undefined.⁵

A concrete example will serve to introduce more of the necessary terms.

A general at a teleconference⁶ writes out a message for transmission. In so doing he functions, from the viewpoint of communications, as a SOURCE. The MESSAGE

³ Much of Fano's first report (see the first paragraph of this review) is devoted to a slow building up of this measure of capacity. He makes it eminently clear why one chooses the logarithm to the base 2. (For readers who have forgotten their high-school arithmetic: the logarithm of n to the base b is a number x , such that $b^x = n$.)

⁴ Because the assignment of a technical meaning to a word which is frequently used as part of our common-sense vocabulary proves constantly embarrassing in more informal discussion. The replacement will be made even in (otherwise) direct quotations from Shannon's text. Similarly 'shannon' will usually thus replace 'bit per second'.

⁵ Wiener's approach is somewhat different, and specifies at least some circumstances under which we can state exactly how much information is conveyed in a given message. See his *Cybernetics*, ch. 3, esp. 75-6. But it is not certain that Wiener is dealing with the same 'information' as Shannon.

⁶ A type of conference, common in military operations, in which participants at widely distant points communicate by teletype.

consists of a linear sequence of SYMBOLS (MESSAGE-UNITS), each one of which is selected from a repertory of 32 possible symbols: the 26 letters of the alphabet (with no distinction between capital and lower case) and six supplementary punctuation marks, one of which is a space.

At the keyboard of the teletype TRANSMITTER, an operator strikes keys in the order required by the message. This TRANSDUCES the message (or ENCODES it) into a SIGNAL, in this case a set of electrical impulses which will travel along a wire until they reach the teletype receiver. The wire, or alternatively a bandwidth of frequencies of electrical impulse used on the wire, constitutes a CHANNEL. Teletype operates in terms of a stock of 32 SIGNAL-UNITS (or SYMBOLS)—different patterns of voltage variation—assigned in a one-to-one way to the 32 letters and punctuation marks allowable in messages so to be transmitted. These signal-units all require the same amount of transmission time. So far as teletype itself is concerned, therefore, a transmission rate of n signal-units per second would imply the possible transmission of $5n$ binitis of information per second, or a capacity of $5n$ shannons—since the logarithm of 32 to the base 2 is 5. For reasons which we shall examine shortly, teletype never attains this maximum.

At the teletype RECEIVER, the incoming signal is retransduced (or DECODED), producing once again a message. This message will show nothing of the general's handwriting, of course, but normally it will be 'literally' the same—that is, it will consist of the same letters and other symbols in the same linear order—as the message produced by the general. The RECOVERED message is then handed to a colonel, let us say, who from the viewpoint of information theory is a DESTINATION OR SINK.

In order for teletype to operate at maximum capacity, it would be necessary for each one of the 32 signal-units to be equally probable at all times during the transmission, regardless of which signal-units had already been transmitted. Now so long as the 32 signal-units are assigned in a one-to-one fashion to English letters and punctuation marks, this condition cannot be met, since all the limitations on sequence of English spelling are directly carried over to the signal. Since the letter-sequences QL, TSR, SSS, and the like, never occur in English spelling, the corresponding sequences of signal-units will never leave the transmitter. Since *CHL (* = 'space') is relatively infrequent, while *CHE is rather more common, the same differences of frequency will appear in the utilization of the various signal-units. After the signal-unit assigned to T, the probability of the one assigned to H will be higher than that of the one assigned to C. All such deviations from constant equiprobability of the signal-units represent inefficiencies—the use of more transmission-time than is necessary for the amount of information to be sent.

Greater efficiency can be attained by changing the code which assigns signal-units to message-units. A first step would be to use signal-units of varying durations (though with the same average duration as before), and to assign the shortest signal-units to the most frequent message-units, the longest to the least frequent. Or instead of considering the message-units one at a time, one could determine the relative frequency of all possible sequences of two message-units, and assign the shortest sequences of two signal-units to the sequences most

frequently used. If one does not care how complicated the transmitter and receiver have to be—if this is a sufficiently trivial consideration relative to the cost of transmission of the signal from transmitter to receiver—then such change of code can be continued until the maximum capacity inherent in teletype ($5n$ shannons) is approached. It is worth noting that more efficient coding of message into signal in general requires a delay at the transmitter, which must collect a number of successive message-units to be encoded all at once, and a similar delay at the receiver.

Most communicative systems involve at least some constraints on sequence; that is, some symbols are not followed by certain others, or are followed by various others with different relative frequencies. To handle this, 'we imagine a number of possible states [of a source or a transmitter] . . . For each state only certain symbols . . . can be transmitted (different subsets for different states). When one of these has been transmitted the state changes to a new state depending both on the old state and the particular symbol transmitted' (8). The matter of relative frequency is easily added to this; by considering that each state is characterized by a set of relative probabilities as to which symbol will next be transmitted and, consequently, which new state will ensue.

We can illustrate with English phonemics. Having begun an utterance by pronouncing a /p/, a speaker of English is in a 'post-/p/' state: he can choose any vowel, or /r, l, y, w/, as next 'symbol', but not, for example, /t/ or /k/. The various possible choices have various relative probabilities. If he chooses /r/, he is then in a 'post-/pr/' state, with new limitations and new probabilities: any vowel, but not, for example, another /r/ or /p/ or /l/. And so on. It is to be noted that the post-/pr/ state is not identical with the post-/r/ state, established when a speaker begins his utterance with /r/.

With this scheme, attention can focus on the whole set of possible interstitial states, instead of on the symbols; the latter can be regarded as elements 'emitted' by the source (or transmitter) as it passes from one state to another. Mathematically the great advantage of this way of viewing the matter is that there is a well-understood set of machinery at hand, the theory of Markoff chains, which is immediately applicable.⁷ Any Markoff chain can be described by a square array of probabilities, the entry in a given row and column being the probability that the state corresponding to the row will be next followed by that corresponding to the column. To facilitate further deductions, some limitations have to be imposed on the variety of Markoff chains allowed; a very general limitation, which both renders further deduction possible and also subsumes a wide variety of cases, is that the chain be ERGODIC: that is, the probabilities must be such that there is no state which can never recur.⁸ This seems to me to correspond to the fundamental (and not always overtly expressed) assumption involved in synchronic linguistic analysis: the assumption that we as analysts, like the speakers

⁷ A good new reference on this is W. Feller, *Introduction to probability theory and its application*, chs. 14-17 (1950).

⁸ More strictly, there is no state which has probability zero of recurrence. Impossibility implies probability zero, but not conversely. Note that the term 'ergodic' is currently used in a variety of more or less closely related senses, of which the present use is one of the simpler.

of a language themselves, can ignore the short-term (hourly, daily, yearly) results of continuous linguistic change, and still get valid results; the extent to which this assumption is false is a measure of the rate of linguistic change.

A source (or a transmitter) which emits its symbols with constant equiprobability is generating information at the maximum rate possible within the limits of the finite repertory of symbols it uses and of the rate at which those symbols are emitted. The actual rate at which a source generates information, on the average, is the ENTROPY of the source; the ratio of this to the theoretical maximum is the RELATIVE ENTROPY.⁹ 'One minus the relative entropy is the REDUNDANCY. The redundancy of ordinary English [writing], not considering statistical structure over greater distances than about eight letters, is roughly 50%. This means that when we write English half of what we write is determined by the structure of the language [i.e. of the language and of the writing system] and half is chosen freely. The figure 50% was found by several independent methods which all gave results in this neighborhood.' One method 'is to delete a certain fraction of the letters from a sample of English text and then let someone attempt to restore them. If they can be restored when 50% are deleted [at random] the redundancy must be greater than 50%' (25-6).¹⁰

Shannon's first major result (towards his aim, summarized in the first paragraph of this section) is the following 'fundamental theorem for a noiseless channel' (28):

Let a source have entropy H [bits per symbol] and a channel have a capacity C [shannons]. Then it is possible to encode the output of the source in such a way as to transmit at the average rate $C/H - \epsilon$ symbols per second over the channel where ϵ is arbitrarily small. It is not possible to transmit at an average rate greater than C/H .

The reader will recall our earlier discussion of the efficiency of teletype and methods of increasing it by modification of code. The theorem establishes the outer limits within which such improvement can be brought about.

But there is a factor, not yet discussed, which sets narrower limits: NOISE. In engineering parlance, noise is anything which operates on the signal, as it travels along the channel, in such a way that the received signals are not always the same as the transmitted ones. To be noise, the effect must be random, and thus only statistically predictable. For if one knew in advance, for example, that

⁹ Entropy can be measured in terms of time or in terms of symbols; the latter is useful in dealing with cases such as writing (or other forms of 'information storage'), where the rate per unit time depends on the rate at which the written symbols are read. If the entropy in terms of symbols is H' , and n symbols per second is the average rate of transmission, emission, or reading, then the entropy in terms of time is nH' .

¹⁰ Some of Shannon's discussion preparatory to this is fascinating, particularly on the subject of successive artificial (statistically controlled) approximations to written English (13-5), which underlie another method of determining the redundancy of written English. Those of us interested in such matters as literary style are particularly apt to enjoy the following paragraph (26): 'Two extremes of redundancy in English prose are represented by Basic English and by James Joyce's book *Finnegans Wake*. The Basic English vocabulary is limited to 850 words and the redundancy is very high. This is reflected in the expansion that occurs when a passage is translated into Basic English. Joyce on the other hand enlarges the vocabulary and is alleged to achieve a compression of semantic content.'

precisely every fifth signal-unit would be distorted in the channel, then it would be easy simply to avoid those moments of distortion, and to transmit the entire message during the noiseless intervals.¹¹

This recalls the necessary indeterminacy (for receiver and destination) in messages themselves, if any information is to be transmitted. If the receiver or destination knows in advance what message is going to be transmitted, its transmission conveys no information; if the receiver knows in advance what distortions are going to be imposed on the signal in the channel, those distortions are not noise and do not interfere with the transmission of the message. In fact, since noise is necessarily random, it is possible to characterize a 'noise source' in precisely the same way that we characterize an information source: a noise source emits an undesired 'message', with a statable entropy, and this undesired 'message' interferes with the reception of the desired one. Put another way, if part of the capacity of a channel is used for the transmission of noise (undesired 'message'), then just that much of the capacity is unavailable for the transmission of the desired message.

Occasional misprints in an article, or errors of transmission in a telegram, do not usually interfere with the intelligibility of the article or telegram for him who reads or receives it. Such misprints or errors of transmission are the result of noise (or, with a slight change of emphasis, can be said to constitute noise). The reason for usual intelligibility despite such noise is perfectly clear: the redundancy of written English. Here, then, is the importance of redundancy: channel noise is never completely eliminable, and redundancy is the weapon with which it can be combatted.

The capacity of a noisy channel is obviously not definable in the same way as that of a noiseless channel. If at the receiver the entropy of the incoming signal is actually equal to the capacity of the channel on the assumption of no noise, then a certain portion of that entropy is in fact due to noise, and only the remainder constitutes the effective maximum capacity of the channel—assuming that at the transmitter the message is being encoded into the signal in the optimum way. This defines the capacity C of a noisy channel (38), and this definition proves to be a generalization of the earlier one for a noiseless channel, in that if zero be assigned to the noise factor in the new definition, the result is the old one.

It may seem surprising that we should define a definite capacity C for a noisy channel since we can never send certain information in such a case. It is clear, however, that by sending the information in a redundant form the probability of errors can be reduced. For example, by repeating the message many times and by a statistical study of the different received versions of the message the probability of errors could be made very small. One would expect, however, that to make this probability of errors approach zero, the redundancy of the encoding must increase indefinitely, and the rate of transmission therefore approach zero. This is by no means true. If it were, there would not be a very well defined capacity, but only a capacity for a given frequency of errors ...; the capacity going down as

¹¹ On page 48 Shannon gives an example of noise which is indeterminate within certain precisely defined (determinate) limits, and of a method of counteracting its effect completely; this is, in a sense, only a more complex example of 'determinate' distortion than that given here.

the error requirements are made more stringent. Actually the capacity C ... has a very definite significance. It is possible to send information at the rate C through the channel *with as small a frequency of errors ... as desired* by proper encoding. This statement is not true for any rate greater than C . If an attempt is made to transmit at a higher rate than C , say $C + R_1$, then there will necessarily be an equivocation [an uncertainty at the receiver as to the correctness of any single received signal-unit] equal to or greater than the excess R_1 . Nature takes payment by requiring just that much uncertainty, so that we are not actually getting any more than C through correctly. (39)

These facts Shannon then rewords more formally as his 'fundamental theorem for a discrete channel with noise'. It turns out that noise can be effectively combatted only in sets of messages of relatively great length; that is, no matter what the coding, sufficiently short messages may always, with a probability greater than zero, be distorted irrecoverably.

We have now followed Shannon's discussion (with many omissions of detail) through the first two of his five numbered sections (through page 48). In the remaining three sections (49-81; 82-91 are appendices) there is a fundamental change of conditions. The systems dealt with so far have all involved sources which emit DISCRETE sequences of message-units, transmitters which send and channels which carry DISCRETE sequences of signal-units. An example is written English, a series of separate symbols, each of which is one or another letter of the alphabet (or punctuation mark). From this point onward Shannon's concern is with CONTINUOUS transmission. An example is the speech signal—the train of sound waves produced by a speaking human being, which have measurable characteristics for any and every value of the continuous independent variable, time. Similarly, a speedometer needle delivers a continuous report on the linear velocity of the vehicle, and there is a continuous band of possible velocities.

From either the engineering or the mathematical point of view it is here that Shannon's argument becomes really interesting (and correspondingly difficult to follow). But for our purposes, we need follow no further. Suffice it to say that in the continuous case, as in the discrete, it proves possible to establish measures for the amount of information, the amount of noise, the entropy of a source, the capacity of a channel, relative entropy, and redundancy, all of which are in general analogous to the corresponding notions worked out for the discrete case.

LINGUISTIC APPLICATIONS

We discuss the linguistic applications of communication theory in six lettered subsections.

A. Continuous and Discrete in Speech

The acoustician examines speech-signals and reports that they are continuous. The linguist examines them and reports that they are discrete. Each uses operationally valid methods, so that both reports must be accepted as valid within the limits defined by the operations used, and the apparent contradiction between the reports constitutes a real, not an imaginary problem.

Neither the acoustician nor the linguist is able to examine the speech-signal

directly, as sound-waves passing through the air. The acoustician currently makes use of oscillographs and spectrographs, both of which transduce the speech signal into a visual report which can be examined at leisure. The transductions involved are quite complex, and do not always give facsimile-type accuracy; but whatever difficulties may be implied by this, at least one thing is certain: oscillographs and spectrographs do not impose a spurious appearance of continuity on a signal that is actually discrete.

The linguist, in phonemicizing, uses no hardware; but he, also, is unable to examine the speech-signal directly. The ear and the associated tracts of the central nervous system constitute a transducer of largely unknown characteristics; in what follows we shall attempt to deduce at least a few of these.

A continuum can be transformed into a discrete sequence by any of various QUANTIZING operations; the notion of quantizing is familiar enough to communications engineers, though the quantizing operations used in electronic communications are all quite arbitrary. Similarly, a discrete sequence can be transformed into a continuum by what might be called a CONTINUIZING operation. Now if the continuum-report of the acoustician and the discrete-report of the linguist are both correct, then there must be, for any given shared body of raw material, a quantizing operation which will convert the acoustician's description of the raw material into that of the linguist, and a continuizing operation which will do the reverse; the desired quantizing and continuizing operations must be inverses of each other.

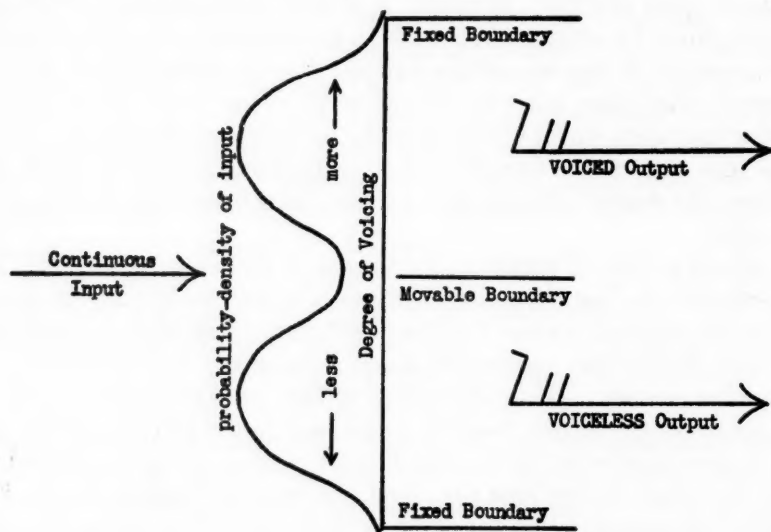
Joos affords a point of departure for the search for these operations:¹² 'Let us agree to neglect the least important features of speech sound [the speech-signal], so that at any moment we can describe it sufficiently well with n measurements, a point in n -dimensional continuous space, n being not only finite but a fairly small number, say six. ... Now the quality of the sound becomes a point which moves continuously in this 6-space, sometimes faster and sometimes slower, so that it spends more or less time in different regions, or visits a certain region more or less often. In the long run, then, we get a probability-density for the presence of the moving point anywhere in the 6-space. This probability-density varies continuously all over the space. Now wherever [one] ... finds a local maximum of probability density,' there the linguist finds an allophone; and 'there will be not only a finite but a fairly small number of such points, say less than a hundred.'

By regarding the moving point as input, and adding certain further specifications, we shall convert Joos's description into that of a transducer. Forgetting for the moment about the probability-density, let us imagine, in the 6-space, a honeycomb of cells, with boundaries between them, such that every point in the space (save those on the boundaries) is in one or another cell. In each cell there is a trigger, which is activated whenever the moving point stays within the boundaries of that cell for a sufficiently long time. Each trigger, when activated, transmits a single output signal, so that the continuous operation of the transducer produces a discrete series of outputs. Finally, imagine that the boundaries

¹² Description of language design, *Journal of the Acoustical Society of America* 22.701-8 (1950). Joos's number 6 is purely arbitrary; one might better replace it throughout by n .

are made of perfectly flexible rubber, so that the location of the different boundaries is not fixed; indeed, at a given time one cell may be enlarged so as to include almost the entire relevant volume of the 6-space, compressing the others to very small size. In a given short interval of time, the output of the system is a function of the input and of the location of the boundaries between the cells. Now we shall specify that the location of the cell boundaries, at any given moment, is a function of the immediately preceding succession of N output signals (NOT input), where N is some fairly large number.

Such a system will indeed map a continuous input into a discrete output. If the details of its construction are based on both the acoustic analysis and the phonemic analysis of the speech-signal of a given language, then the system will transduce the acoustician's description into the linguist's description, or, what amounts to the same thing, will transduce the speech-signal in the physical sense into a linear sequence of allophones in the linguist's sense.



A one-dimensional reduction will serve as illustration. Suppose for a moment that the only relevant feature of speech is voicing versus voicelessness. Even for just this one feature the acoustician tells us that the voice-bar on a spectrogram may have virtually an infinity of degrees of prominence (or could have on an ideal spectrogram), while the linguist knows that, in French or English for example, there are two and only two degrees of contrast along this particular scale. In the appended diagram the continuous input arrives from the left. The curve represents the probability-density along this one dimension; our reason for giving the curve its particular shape will appear later. At the top and bottom of the scale are two fixed boundaries. Between them is a movable boundary. If a sufficiently long stretch of input falls above this movable boundary, the top trigger is activated and the voiced output signal is transmitted; similarly below the movable boundary. At a given moment, if the probability of voiced as the next signal is extremely high, then the movable boundary will slide very far down the scale,

so that almost any next input-stretch will activate the voiced trigger rather than the voiceless trigger; and vice versa. The output will then be a discrete succession of signals, each of them either voiced or voiceless.

Let us return to the way in which the location of the movable boundaries is determined. It is determined, in the first instance, by the preceding N output signals. Since this preceding output depends, in turn, on input, one might suspect that the preceding N output signals, as a factor conditioning what happens to a given bit of input, could eventually be bypassed, as we consider the progressively earlier history of the transducer. But each of these preceding N output signals was also dependent, because of the principle which governs boundary locations, on the N output signals which preceded it, so that there can be no hope of bypassing the conditioning effect of 'previous output' save by tracing the operation of the system back to some initial state, before any input at all. So far, such an initial state is undefined, but we will return to it in a moment. Examining the system in current operation, what the acoustician is able to do is to describe probability-densities. What the linguist can do is to state, on the one hand, the topological structure of the honeycomb of cells, and on the other, the contingent probabilities of each output signal after each sequence of N preceding output signals. Part of this is what the linguist currently does when he determines and describes phonemic structure; the remainder is what he could do with no further technique but some simple statistics. What the acoustician and the linguist working together can do is to determine the metric structure of the honeycomb under various conditions, including the relation of boundaries to local maxima of probability-density.

Our assumption is that, in a way, the above description of a transducer and a type of transduction applies to the processing to which the incoming speech-signal, after it impinges on the ear of a hearer, is subjected, within the central nervous system of the hearer. Because the linguist has such a transducer within himself, which operates on the speech-signal before he can begin to examine and analyze it, he cannot (using no hardware) perceive the continuous nature of the speech-signal, but necessarily interprets it as discrete.

It is known in information theory that if a continuous signal is quantized, no transduction is available which will recover the original continuous signal exactly.¹³ Our assumption is that this is irrelevant, since the linguistic information carried by the speech signal accounts for only a small part of the total entropy thereof; the quantizing transduction performed within the hearer need only recover as much from the speech signal as was put into it by the speaker. The act of speaking is also a transduction; it converts an inner discrete flow of allophones, inside the speaker, into the continuous speech signal. The apparent discrepancy between the acoustician's report and that of the linguist is then due to the fact that they are tapping the same complex communications system at different points: the acoustician taps it in the external channel, where the in-

¹³ Shannon's Theorem 13 (p. 53) does not contradict this assertion. It allows us to transduce a band-limited continuous signal into a discrete signal, but requires for complete recoverability the use of an infinite repertory of discrete signal-units. All that can be done with a finite repertory is to approximate complete recoverability as closely as it is desired.

formation indeed flows in a continuous form; the linguist taps it after the quantizing transduction of hearing, or before the continuizing transduction of speaking, or both. Edward Sapir characterized these two transductions beautifully, though in a different terminology, when he wrote in 1921: 'In watching my Nootka informant write his language, I often had the curious feeling that he was transcribing an ideal flow of phonetic elements which he heard, inadequately from a purely objective standpoint, as the intention of the actual rumble of speech.'¹⁴

Since the linguist does not investigate these postulated 'inner flows' with scalpel and electroencephalograph, it is proper to ask just what procedure he does use. To answer this we ask what might be meant by our earlier reference to an 'initial state' of a quantizing transducer. If our requisite quantizing and continuizing transducers exist in human nervous systems, then the 'initial state' of either is the state to be found in a child before he has begun to learn his native language. As he learns his native language, the child has access to only certain kinds of evidence to decide whether two stretches of incoming speech-signal are instances of 'the same' signal (that is, are phonemically the same) or not: the physical properties of that speech-signal, insofar as human physiology can measure them, and the conditions under which the speech-signal arrives—in short, physical similarity and similarity of meaning. The building of the quantizing transducer in the child proceeds by trial and error, errors being corrected by failure to adapt appropriately, as with any learning process.

If the linguist cannot open up the transducers in the head of a native speaker of a language, he can do something just as effective: he can himself become as a little child (insofar as the specific language is concerned), and build within his own nervous system a reasonable facsimile of the transducer that the native speaker has. A reasonable facsimile is enough, for if a language is spoken by more than one person, the transducer inside any one person is no more than a reasonable facsimile of that inside another. This is what the linguist can do, and does. In addition, he does something that the child cannot do: he keeps an overt record of the structure of the system which is being built internally. This overt record in due time becomes his description of the phonemic system which has come to exist inside his own nervous system.

We can now see why local maxima of probability-density, determinable by acoustics without the help of the linguist, will show some correlation with the linguist's allophones. If there were no such correlation, the process of learning a language, and equally the linguist's problem of phonemicizing a language, would be impossible. That is why the curve of probability-density, in the simplified one-dimensional case of the diagram, was drawn with two local maxima.

The above considerations, if valid, define the problem of acoustic phonetics: accepting the phonemic findings of the linguist and the acoustical findings of spectrograms or oscillograms, it is the task of acoustic phonetics to determine and describe the very complicated nature of the two transductions which relate the two sets of findings.

¹⁴ *Language* 58 fn. 16 (1921).

B. Phonemicizing

When the communications engineer deals with telegraphy, teletype, or the like, the signal-units which are involved are either presented to him as a fait accompli, or are invented by him, within certain constraints, to suit his own fancy. When the linguist begins to examine a language, the signal-units used in that language (speaking phonemically) are neither presented to him for effortless taking, nor can he invent them; he has to go through certain operations to DISCOVER them.

The nature of these operations was in part expounded in section A. Generally, however, the linguist does not stop with allophones; he proceeds to group allophones into phonemes. It is interesting to note that the criteria by which this is done, and the criteria which lead to the choice of one possible phonemicization over another, reflect a consideration which can be stated in terms of information theory: he prefers the phonemicization which yields maximum average entropy per signal-unit. The entropy of speech, in terms of whatever discrete signal-units may be discovered or set up by the analyst, is presumably invariant from one possible valid phonemicization to another, varying only with the rate of articulation. But the ENTROPY PER SIGNAL-UNIT is not thus invariant. For example, by identifying the unaspirated stops of English *spill*, *still*, *skill* with the aspirated stops of *pill*, *till*, *kill* (or, equally well in this connection, with the voiced stops of *bill*, *dill*, *gill*), we obtain a smaller repertory of signal-units with greater freedom of occurrence relative to each other, instead of a greater number with less freedom. This directly implies greater average entropy per signal-unit. Perhaps Harris has carried phonemic procedure further along this line than anyone else.¹⁵

Now a linguist, in the course of phonemicizing, manipulates his data not in the form of actual allophones, but in terms of graphic symbols in some way assigned to them. This is, of course, quite unavoidable; but it implies that a further transduction has been applied to the data before it is in manipulable form, and errors can always creep in during such a transduction. Furthermore, one of the chief practical aims of phonemic study is the devising of orthographies, at least as a basis for analysis of the language on other levels, if not for purposes of literacy. As a result, our current phonemic doctrines contain, to varying degrees, criteria based on orthographic rather than on purely linguistic considerations. This admixture is found in its most extreme form in Pike,¹⁶ but is by no means absent elsewhere.

It is difficult to explain to a communications engineer what we mean by phonemicizing. The following explanation has worked in a few cases; it is also helpful because it shows more clearly how we can draw the line of demarcation between purely linguistic considerations and extraneous orthographic or notational considerations.

Suppose (goes the explanation) we devise a set of messages, each one of which consists of a sequence of ten digits, every digit being either '1' or '2', and all

¹⁵ *Methods in structural linguistics*, chs. 1-11 passim (1951).

¹⁶ *Phonemics* (1947); Grammatical prerequisites to phonemic analysis, *Word* 3.155-72 (1947).

possible sequences of this kind being allowable in the system. Given equiprobability, the system is capable of transmitting on an average 10 bits of information per message. Call this binary code (a).

Now suppose that we decide, quite arbitrarily, to replace a random selection of the occurrences of the symbol '1' in these messages by the symbol 'r'. The message which in code (a) would appear always as '1211212221' will now be transmitted in that same shape, or alternatively in the shape 'r211212221', '12r1212221', 'r2112r2221', and so on. This modification gives us code (b). Nothing essential in code (a) has been changed; the entropy per message is the same as before; but we have a case of what the linguist calls 'free alternation'. That is, '1' and 'r' are physically distinct ('phonetically different', different 'allophones'), but do not contrast. Yet if one began the study of a system which used code (b) by examining sample messages in it, one's first impression would be that the repertory contained three different symbols rather than two. Only statistical analysis would show that the code was essentially binary rather than ternary. The linguist would conclude, in due time, that '1' and 'r', though 'phonetically' different, were 'phonemically' the same: in free alternation, and 'phonetically similar' in that the shapes '1' and 'r' resemble each other more than either of them resembles '2'.

Code (b') is formed from code (b) by using the shape '3' instead of the shape 'r'; the sample message given above might now be transmitted as '3211232223', '1213212221', and so on, all equivalent. With respect to information, this is still an inessential change, and the code is fundamentally binary. But the linguist at this point would balk at phonemically identifying '1' and '3', despite the fact that they are in free alternation, since the shapes '1' and '3' do not resemble each other any more than either of them resembles the shape '2'. The factor of 'phonetic similarity' is lacking.

Next, suppose we keep three physically distinct symbols, '1', '2', and 'r' for code (c), '1', '2', and '3' for code (c'), but set up an instance of complementary distribution instead of free alternation. Starting with the messages of code (a), we replace '1' respectively by 'r' and '3' wherever in code (a) '1' is immediately followed by '2', but otherwise keep the symbol-shape '1'. Once again, there is no essential change in the code from an information point of view. But in code (c) the linguist will note the complementary distribution and the phonetic similarity, and will call '1' and 'r' allophones of a single phoneme; while in code (c') the presence of the former and absence of the latter will lead him to set up three phonemes, despite the lack of contrast between two of them.

Finally, we devise codes (d) and (d') by a rather different step, based on messages as they appear in codes (c) and (c') respectively: we delete from the messages all occurrences of 'r' or '3' which are immediately preceded by the shape '1', but retain those occurrences of 'r' or '3' which are initial in a message or which are immediately preceded by '2'. For example:

code (a)	1221221112	1121211121
code (c)	1r2r2211r2	1r2r211r21
code (c')	3223221132	3323211321
code (d)	122r22112	12r21121
code (d')	322322112	12321121

The messages of code (d) or (d') stand in one-to-one relation to those in code (c) or (c'), and hence to those in code (a). The latter can be recovered from those in (c) or (c') by operating on them as follows: wherever the sequence '12' occurs, insert an 'r' (or a '3') between the '1' and the '2'; then replace all instances of 'r' (or '3') by '1'. In (d) and (d') the ENTROPY PER MESSAGE is the same as in code (a).

However, neither the information theorist nor the linguist would claim that codes (d) and (d') are identifiable with code (a) in quite the same way as codes (b) and (c).

In (d) and (d'), the ENTROPY PER SYMBOL is different from that in (a); one has here a real ternary code, in which there are considerable restrictions in sequence, restrictions which reduce the overall entropy of the code to that of a binary code of type (a). From the linguistic point of view, 'r' and '1' are in contrast in (d), and '3' and '1' in (d'), because there are such initial sequences as 12... versus r2... (or 32...), and such medial sequences as ...212... versus ...2r2... (or ...232...).

Long before we get this far, the communications engineers will have asked us: Why bother to make such replacements and introduce such complications? Why not stick to just one shape for each symbol in the repertory? The answer to this is clear from the first paragraph of this subsection: such complications, arbitrary though they may be, are of the sort encountered by the linguist as he examines languages; they are not invented by him for the pleasure of complexity, but simply have to be accepted. When the linguist works with a language, what he finds accessible to study is in the first instance a body of messages exhibiting a large variety of such complications and limitations of sequence. His task, as phonemicist, is essentially to determine the precise nature of the allophonic units, the nature of the interrelations between them (free alternation, complementation, contrast, and the like), and in these terms to ascertain the various statistical restraints on sequence. So much, and only so much, is his LINGUISTIC task. He may do this in terms of allophones (cells in the 6-space described above), or he may break allophones down into simultaneous components and regard them as his fundamental units (equivalent to taking the coordinates of the local maxima of probability-density in the 6-space and handling those in each dimension as separate units).¹⁷

Essentially irrelevant orthographic considerations come in as soon as the linguist proceeds to the point of phonemic identification. When, in code (b) or (c), the linguist says that '1' and 'r' are 'the same phoneme', he is simply summarizing the following facts: '1' and 'r' are, within the system, phonetically similar; they do not contrast; messages in this code can therefore be transduced, in a manner that preserves the entropy, into a code with one fewer symbol; in devising a writing-system one can make use of this fact and eliminate a symbol needed earlier. When, for code (b') or (c'), the linguist refuses to make a similar statement, he is reflecting an aspect of the meaning of the term 'phoneme' which

¹⁷ As described in the parenthesis, the result would be acoustic componential analysis. What we normally do is to change coordinates in such a way as to get articulatory components. The properties of the system are in theory invariant under any such change of coordinates, no matter how complex; if in practice they are not, it is because we do not yet understand well enough how to change the coordinates. This is part of the task of acoustic phonetics.

is irrelevant with respect to information—the requirement of phonetic similarity—and is therefore simply choosing a different terminology to report the remaining facts of the case. Some linguists would be tempted to ‘phonemicize’ code (d) by saying that where overtly one has the allophonic sequence ‘12’ there is ‘really’ a variety—a zero alternant—of ‘1’ between them; to base a writing-system on this consideration would clearly be feasible in an unambiguous way, but within phonemics such a step is not valid.

The communications engineer is right in not understanding fully what linguists mean by phonemics, for we linguists have been fairly muddled in our phonemic thinking. The establishment of phonemic units can be rendered relatively non-arbitrary by accepting the criteria of phonetic similarity and of contrast versus no contrast, and by preferring that otherwise valid phonemicization which maximizes average entropy per symbol. But the selection of these criteria is itself arbitrary. A redefinition of the aims and procedures of phonemic analysis along the lines suggested above, and a clearer segregation of purely orthographic considerations, is a desideratum.

C. The Entropy and Redundancy of Speech

Speech, examined linguistically, is discrete, but it is not necessarily linear. In telegraphy, which is linear, if two signal-units are to be transmitted, one of them can be sent either before or after the other, but there is no third alternative. In speech there is a third possibility: the signal-units may be simultaneous. This arrangement is possible, of course, only for certain phonemic units relative to certain others, not freely for all, even if we go to the extreme of componential analysis and say that at practically any time more than one phonemic component is being transmitted. Nevertheless, this greatly complicates the problem of measuring the entropy of speech. The mathematical frame of reference worked out by Shannon for such measurements can be applied only if we interpret all simultaneous bundles of components, or of short components and temporal segments of long components, as distinct phonological units, a procedure which does indeed portray speech, phonologically, as linear. In English, for example, /á/ (the vowel with loud stress), /â/, /ã/, and /a/ would thus have to be interpreted as four distinct units, rather than as one vowel and four different accompanying and simultaneous stresses; and this set of four would have to be multiplied by four again to account for differences in phonemic tone. Such a procedure is obviously highly inefficient for most linguistic purposes.

An alternative is to modify Shannon’s mathematical machinery so as to take care of a set of several linear sequences of symbols transmitted in parallel, where there are statistical restraints not only among those in the same linear sequence, but also between those in one linear sequence and those in others. I have no idea how complicated the necessary mathematical machinery might be, but I suspect that it would be very complicated indeed.

In the face of these difficulties, it may seem absurd to give any figure at all for the entropy of speech; we shall nevertheless state that the entropy of English, at normal conversational speed, seems to be very roughly in the neighborhood of 50 shannons. This figure may be off by as much as one hundred per cent, and

is more likely to be an overestimate than an underestimate. For our immediate purpose this rather gross inaccuracy does not count for much, since we want to compare the entropy of speech, analyzed phonemically, with the capacity of the channel used by the speech signal. This channel is a bandwidth of acoustic frequencies; if fully utilized, it could carry 50,000 shannons.¹⁸

The discrepancy is astonishing. Neglecting noise, it would imply a relative entropy of the source of only 0.1%, a redundancy of 99.9%. This would reveal human speech as one of the least efficient communicative systems in existence.

But there are other factors in the situation which render it a bit less striking. A speech signal carries more information than just that imposed on it by the phonological structure of what the speaker is saying. Some of this information serves to identify the speaker, since we do manage somehow to tell people apart by their voices. Some of it tells the hearer about the mood or state of health of the speaker—whether he is angry or contented, whether or not he has a spring cold.¹⁹ Of course, linguistically relevant portions of the speech signal may also carry information, indirectly, about all of these matters, but that is part of the 0.1% and does not concern us. In the normal face-to-face situation of speech communication, there is a good deal of interchange of information which is not carried by the speech-signals at all, but by the continuous train of socially conditioned bodily movement and gesture which both accompanies speech and goes on during silence.²⁰ If we could measure the capacity of this channel—for certainly it is one—and add that to the outside capacity of the channel of the speech signal, the relative figures would be even greater.

No one knows how much of the capacity of the speech signal is taken up by metalinguistic and speaker-identification information, but it may be a good deal. It is for all these reasons that the linguist has very little useful advice for telephone and radio engineers. Their job is to deliver the whole speech signal, with as high a fidelity as the public wants, or as high as is possible in terms of the cost the public will put up with. Measurement of fidelity has to be made psycho-acoustically in terms of the whole speech-signal, not just in terms of its linguistic content; it may be important to be able to tell over the telephone that someone has a cold in the head.

Furthermore, language sometimes must operate under extremely noisy conditions. The high linguistically relevant redundancy of the speech signal can be interpreted not as a sign of low efficiency, but as an indication of tremendous flexibility of the system to accommodate to the widest imaginable variety of noise conditions. And here we mean not only 'channel noise', which is noise in Shannon's sense, but 'semantic noise', discrepancies between the codes used by trans-

¹⁸ R. M. Fano, The information theory point of view in speech communication, *Jour. Acoust. Soc. Am.* 22.691-6, esp. 694 (1950).

¹⁹ A certain proportion of the articulatory and auditory but nonlinguistic milieu of linguistic signalling is highly organized and culturally transmitted. This portion constitutes what G. L. Trager and H. L. Smith Jr. take as the object of METALINGUISTIC study; see now Smith, *An outline of metalinguistic analysis* (1952).

²⁰ The significance of gesture has traditionally been underestimated, and its 'naturalness'—the extent to which various cultures agree on gestures and their meanings—has been vastly overestimated. See R. I. Birdwhistell, *Introduction to kinesics* (1952).

mitter and receiver, the kind of noise despite which we often understand someone with a speech-pattern widely different from our own.²¹

It is worth while also to consider the ratio of the amount of information which can be carried by any one phonemic contrast in a language, given the statistical structure of the whole phonemic system, to the total entropy of the phonemic system. Different contrasts obviously carry different amounts of 'functional load'; just as obviously, no single contrast ever carries any very high proportion of the whole load. The redundancy of a phonemic system is so high that most of the time a hearer need receive accurately only a small percentage of the phonemic units transmitted by the speaker, in order to reconstruct the whole message. This bears on the problem of phonetic change. Any single contrast in a phonemic system can be lost, by phonetic change, without the speakers' being any the wiser. This also militates against Hoenigswald's theory that coalescence of phonemes can be brought about only by dialect borrowing.²²

D. Phonology and Tactics

When the linguist goes beyond the phonological level to the tactical (or 'grammatical' or 'morphemic') level, he finds another way in which to regard utterances as discrete messages, the units in this case being not phonemes but morphemes. It is by no means certain that calculation of the entropy of speech in terms of morphemes will give the same results as calculation in terms of phonemes—though it is certain that the computation is vastly more difficult.

There is a way to approach the relationship between tactical pattern and phonemic pattern which may be physically and physiologically meaningless (though not necessarily so), but which nevertheless has some utility. This is to say that, just as the external speech-signal represents a continuizing transduction of an internal discrete phoneme flow, so this phoneme flow itself represents a transduction of an even deeper morpheme flow. The morphemes of a language, in this portrayal, cease to be CLASSES of morphs, and become rather message-units on this deeper level which are REPRESENTED by morphs on the phonemic level. The morphophonemics of a language is then a set of rules for transducing morpheme-sequences to phoneme-sequences. And in the hearer, after the more superficial quantizing of the arriving speech-signal into a discrete phoneme flow, the morphophonemic principles of the language are applied backwards to the phoneme flow to recover a morpheme flow.

To make this concrete, let us image a unit called a tactics box in the brain. This tactics box passes through a series of states. Each passage from one state to another is accompanied by the emission of a morpheme, and the new state depends both on the old state and on the morpheme emitted, as well as on two additional factors to be mentioned presently. When the box is in any given

²¹ Semantic noise is discussed very briefly by Weaver towards the end of his paper in the volume under review. The reviewer's paper *An approach to the quantification of semantic noise*, *Philosophy of science* 19.257-60 (1952), though set up in terms of a highly oversimplified model, perhaps shows how communication can take place despite failure to agree completely on code-conventions.

²² See for example his review of Hall's *Leave your language alone*, *Classical weekly* 42.250 (1949).

state, there are various probabilities that it will pass next to each of the various other states, and thus the same probabilities that it will emit next each of the various morphemes which constitute its repertory. Insofar as these probabilities are determined by previously emitted morphemes, they constitute the tactical structure of the language. The emitted stream of morphemes gets encoded into a stream of phonemes; there is delay at the transducer which does this, since the proper stream of phonemes for a given string of morphemes often depends on several successive morphemes, not just on one (*wife* is encoded into /wayv/ if the next morpheme is the noun-plural -s, and -s is encoded into /z/, rather than /s/, when the preceding morpheme is *wife*). The stream of phonemes is smeared by articulation into a speech signal; this, entering someone else's ear, is quantized again into a stream of phonemes, and then decoded into a stream of morphemes, which is fed into the tactics box of the hearer. For a tactics box is a combined source and sink; the impact of incoming morphemes is a third factor conditioning the sequence in which the box passes from one state to another. We can add the specification that on some occasions emitted morphemes are shunted directly back into the emitting box, or are converted into phonemes and then decoded back into morphemes in the same brain, instead of breaking all the way out in the form of overt speech; this is 'thinking in words'.

The last factor that conditions the probabilities of change of state in the tactics box is all that goes on in the rest of the central nervous system: the constant feeding in of a stream of perceptions, the retention of some of these, the reorganizing of others, the delaying of still others. We can say that the conditioning of the tactics box by this factor is the SEMANTICS of the language.

Since there is currently no way in which all this can be disproved, it does not qualify as a scientific hypothesis; it is merely a terminology. It should be apparent, however, that as a terminology it affords us a way to bring existing techniques in linguistics and existing methods in communication theory jointly to bear on the workings of language.

E. Immediate Constituents

Usually in tactical analysis (less often in phonological) linguists make use of the procedure of immediate constituents. Sometimes this is regarded merely as a way of attaining descriptive statements in economical form; sometimes it is regarded as an objective factor in linguistic structure which must be ascertained in the course of analysis. In either view, there would appear at first sight to be nothing in information theory to parallel it. As we shall show, however, there is.

At various points in an utterance in the course of transmission, say at the ends of successive morphemes, the degree of indeterminacy as to what may come next can in theory be computed. The indeterminacy is greater if the number of possible next morphemes is greater; with a given number, the indeterminacy is greater if the probabilities are nearly equal, less if they diverge widely from equality. Now generally in current practice, and universally in a theoretically possible optimum procedure, a linguist makes his primary cut of a composite form at that point where the indeterminacy is greatest. The form *red hats* (we ignore suprasegmental features) consists of three morphemes, with a cut between *red* and *hat*

and another between *hat* and *-s*. It would seem that the indeterminacy after *red* is greater than that after *red hat*; certainly it is if we regard the singular form *hat* as involving a morpheme of zero shape 'singular', since in that case there is only a small handful of morphemes which can immediately follow *red hat*: this singular morpheme, or *-s*, or *-ed* (*red-hatted*), or perhaps one or two others. Some forms, it may be, are not actually so cut, because the pressure of their similarity to large numbers of forms where the cutting is unambiguous may lead us to go counter to this first and fundamental criterion. Thus there is probably greater indeterminacy after the *hermetic*- of *hermetically sealed* than there is after the whole first word, but we would all choose to cut first between the words.

Shannon has conducted experiments in ordinary English orthography,²³ and the reviewer has conducted similar ones, with the proper audiences, in terms of phonemic notation, the results of which bear on the stated correlation between IC-analysis and information theory. One decides on a sentence which is totally unknown to the audience, and writes it down. Then one has the audience guess—without any hints—the first letter (or phonemic symbol) of the sentence, and records the number of guesses made, up to and including the right guess. Then the second letter (or phonemic symbol) is guessed, and the third; spaces in the orthographic form, and open junctures in transcription, count as symbols and have to be guessed along with the others. As might be imagined, the number of guesses necessary for each successive symbol varies considerably; this number decreases sharply within the body of a word or morpheme, and in general increases when any word boundary or morpheme boundary is reached. And one can discern some tendency for a larger number of guesses to be required at some cuts between morphemes or words than at others; the greater number usually correlates with a more elementary cut between ICs.

F. Writing

When one writes, one transduces a linguistic message into a different form; reading is the inverse transduction. In some writing systems (Chinese) the linguistic message is taken in morphemic shape for such transduction; in others (Finnish), it is taken in phonemic shape; in most—including actually both of the extremes named—both elements are involved in various proportions.²⁴

Most traditional writing systems provide no machinery for the indication of certain relevant features of speech. In English writing, for example, there are no conventions for indicating stresses and intonations. The information carried in speech by stresses and intonations is therefore either lost in the transduction to writing, or is carried by the use of more morphemes of other types.

A short count seems to indicate that about one-sixth of the morphemes of English are lost in the act of writing in just this way. This was determined by counting the number of morphemes graphically indicated in several passages of written English, then counting the number which occurred when the same

²³ Shannon, Prediction and entropy of printed English, *Bell System technical journal* 30.50-65 (1951).

²⁴ This is worked out in more detail in the reviewer's paper for the Third Annual Conference on Linguistics and Language Teaching (Georgetown University), to be published soon.

passages were read aloud. The specific morphemes added in reading aloud may not match the ones spoken by the writer preparatory to writing, but certainly the number of occurrences of such morphemes must be about the same.

The fact that written English is intelligible despite this loss, and that it can often be read aloud in a way which restores the lost morphemes with reasonable accuracy, implies, of course, a degree of redundancy in spoken English that we already know to be there. Nevertheless, not all passages of written English are intelligible. Given a writing-system which forces such loss, a good WRITING STYLE is definable as one which compensates for the loss by the use of more segmental morphemes. All of us have seen, particularly in newspapers, passages which were not in good writing style.

The devising of utterances in a good writing style to be transduced into written form finds its analog in many other communicative situations. In preparing a written message for telegraphic transmission, for example, we refrain from any significant use of the contrast between lower-case and capital letters, since that distinction is not maintained in the transduction to telegraphic signal. The general at the teleconference of our first section imposed the same restraint on what he wrote, for the same reason. When it is necessary to transmit via telegraph or teletype a chemical formula or a mathematical equation, the task cannot be done by direct transduction; instead, we do the same thing that must be done in sending information about a building-plan or the like: we phrase and transmit a DESCRIPTION of the formula, equation, or building-plan, on the basis of which the destination can construct a more or less reasonable facsimile. Compare with this the difference between a dialog in a novel and on the stage: the novelist writes, *'You can't do that!' he rasped*; the actor on the stage omits the last two words but rasps as he delivers the rest of it.

GENERAL IMPLICATIONS

We have demonstrated that a certain number of problems of concern to linguists can be phrased in the terminology of information theory. We have not proved that such rephrasing makes for added clarity, or leads to the posing of relevant new questions. It is always challenging to discover that a systematization will subsume a larger variety of phenomena than those for which it was in the first instance devised; the discussion in the last section implies not only that information theory is applicable to language, but also that linguistic theory is applicable, with certain restrictions or modifications, to other varieties of communicative behavior. But in our exhilaration we must always guard against misleading analogy and the invalid identification of concepts because the words for them are identical. Otherwise we may think we have obtained genuinely new knowledge when all we have done is to formulate old knowledge in a new terminology.

'Information' and 'meaning' must not be confused. Meaning might be said, in a sense, to be what information is about. For the relatively simple communicative systems that Shannon handles, it is easy to introduce a realistic definition of the meaning of a signal or a portion of a signal: the meaning of a stretch of signal is the stretch of message which was transduced into it and into which, at

the receiver (barring noise), it will be retransduced. In telegraphy the meaning of two dots is the letter I. The meaning of the English written word 'man' is the spoken form *man*. The meaning of the speech-signal produced when someone says *man* is the sequence of phonemic units which compose that word linguistically. The meaning of the sequence of phonemic units is a certain morpheme. But if we inquire into the meaning of the morpheme, information theory cannot help us. Information theory does not deal with the way in which a source maps a noncommunicative stimulus into communicative output, nor with the way in which a sink maps communicative input into a noncommunicative response. Precisely this is the fundamental problem of semantics, and on this score the speculations of linguists such as Bloomfield or of psychologists such as Weiss have much more to tell us than information theory. It is possible that these speculations could afford the basis for an expansion of information theory in valuable directions.

There is a partial analogy between information and energy, which extends also, as a matter of fact, to money; the tendency is strong to make more of this analogy than is justified. Energy-flow is POWER; information-flow is ENTROPY; money-flow (at least in one direction) is INCOME. Energy is measured in ergs or watt-seconds or kilowatt-hours, power in watts or kilowatts; information is measured in binitis, entropy in shannons; money is measured, say, in dollars, income in dollars-per-month. In all three cases it is, in a sense, the rate of flow that counts; energy, information, and money are substantialized (the last actually in the form of pieces of metal or paper, the other two only in words) primarily because we find it easier to think about problems in that way—perhaps because we think in languages of the type that Whorf called Standard Average European.

But there is a law of conservation of energy, while there is no law of conservation of information (I cannot speak for money). At this point the parallelism breaks down. Proof is easy. To supply one hundred-watt light bulb, a generator must transmit one hundred watts of power, plus a bit more to make up for line-loss. To supply ten such bulbs, the generator must transmit ten times as much power. If all the bulbs are turned off, the generator is forced to cease transmitting power—either it is turned off also, or it burns itself out. To supply a receiver with one hundred shannons, a source must transmit information at that rate, plus enough more to counteract noise. But to supply ten receivers with one hundred shannons each, the source need not increase its entropy at all (unless the hook-up produces more noise to counteract). The entire output, minus that lost through noise, reaches each receiver. And if all receivers are turned off, the source can continue to produce and transmit at the same rate. The information in this case does not dissipate (as energy might in the form of heat); it simply disappears. We have all had the experience of continuing to talk over the telephone after the person at the other end of the line has hung up.²⁵

This defect in the analogy has proved uncomfortable for some investigators, who have also, perhaps, been misled by the homophony of 'information' as a tech-

²⁵ The contrast between energy and information appears in biological study: the physiologist is concerned primarily with energy-flow, the psychologist (even the physiological psychologist) primarily with information-flow.

nical term and 'information' in everyday speech. One writer tries to differentiate between ABSOLUTE INFORMATION, 'which exists as soon as one person has it, and should be counted as the same given amount of information, whether it is known to one man or to millions', and DISTRIBUTED INFORMATION, 'defined as the product of the amount of absolute information and the number of people who share that information'.²⁶ Whatever validity there may be in the distinction, neither of the items distinguished can be identified with Shannon's information, and Shannon's work affords no method for quantifying either. If it is necessary to maintain some analogy between an information-system and a power-system, then entropy can better be compared to voltage, since a single generator can supply current at a specified voltage to any number of outlets.

It helps in avoiding this particular error to note that a transducer is a kind of complicated trigger. When a marksman aims and fires a pistol, the energy that he must expend bears no necessary relation to the amount of energy produced by the burning powder in expelling the bullet. When the operator depresses a key on a teletype transmitter, the energy that he uses is not the energy which is transmitted along the wire, and bears no necessary quantitative relation to this. The human brain operates on a power-level of five watts; such a brain guides a derrick which lifts a stone weighing tons. In micrurgy, on the other hand, the operator expends much more energy on his apparatus than it expends on its object.

The distinction between trigger action and direct action is of fundamental importance at least in human life, possibly in the entire physical universe, and one line of argument implies that 'communication' is ultimately definable only in terms of trigger action. In the world of man, the trigger effect of language is too obvious to need discussion. Human artifacts can be classed into tools, which involve no triggers or transducers, and machines, which do. Some artifacts are used for SENSORY PROSTHESIS: telescopes and microscopes are tools for sensory prosthesis, while radar, electron-microscopes, Geiger counters, and the like are machines. Other artifacts are used for MOTOR PROSTHESIS: spades, shovels, wrenches, bows and arrows are all tools, while steam shovels, firearms, and draft animals are machines. Still other artifacts are used for COMMUNICATIVE PROSTHESIS: language itself (or rather, the vibrating air which is the artifaction produced by speech) is a tool, while writing, smoke-signals, and electronic apparatus of various kinds (including mechanical computers) are all machines. This ties in with White's notion of measuring human evolution in terms of the increasing amounts of energy controlled and utilized by human beings.²⁷ It is clear that human tools developed before human machines, and the simpler machines before the more complex. A very late development is the practice of coupling a device for sensory prosthesis, directly or through one for communicative prosthesis, to a device for motor prosthesis, so as to produce an apparatus which will perform as it is supposed to perform without human participation—e.g.

²⁶ L. Brillouin, Thermodynamics and information theory, *American scientist* 38.594-9, esp. 595 (1950).

²⁷ A. L. White, Energy and the evolution of culture, *American anthropologist* 45.335-56 (1943). White gives credit in turn to Morgan.

an electric refrigerator or a radar-controlled anti-aircraft battery. On the level of human understanding, much of man's progress has consisted of a slow clarification as to what can be triggered, and by what means. If you can trigger a fire into cooking your meat, why can't you trigger the sky into giving you rain? The difference between the rites of a shaman and the seeding of clouds with dry ice is that the latter sometimes triggers a release of rain, whereas the former never does. All of this has an important inverse: a deeper understanding of man's role in the evolution of the universe.

The argument which attributes such cosmic significance to communication is based on an identification which may be as fallacious as that of Shannon's 'information' and the common-vocabulary term 'information'. This is the assumption that the entropy of communication theory is physically the same thing as the entropy of thermodynamics. The latter is a measure of the degree of randomness of energy-distribution in a closed physical system; the second law of thermodynamics states that in any such system the entropy increases, by and large, until it is maximum. If information-theory entropy is actually the same thing, rather than something else which can be handled by the same mathematical machinery, then the transfer of information from a physical system represents a local decrease in entropy—an increase in orderliness and pattern. Since the only completely closed physical system in the universe is the universe itself, local decreases in entropy do not controvert the second law of thermodynamics. It is none the less valuable to study the mechanisms by which local and temporary decreases are brought about. With Wiener's elegant discussion of this²⁸ we close our review:

A very important idea in statistical mechanics is that of the Maxwell demon. Let us suppose a gas in which the particles are moving around with the distribution of velocities in statistical equilibrium for a given temperature. For a perfect gas, this is the Maxwell distribution. Let this gas be contained in a rigid container with a wall across it, containing an opening spanned by a small gate, operated by a gatekeeper, either an anthropomorphic demon or a minute mechanism. When a particle of more than average velocity approaches the gate from compartment A or a particle of less than average velocity approaches the gate from compartment B, the gatekeeper opens the gate, and the particle passes through; but when a particle of less than average velocity approaches from compartment A or a particle of greater than average velocity approaches from compartment B, the gate is closed. In this way, the concentration of particles of high velocity is increased in compartment B and is decreased in compartment A. This produces an apparent decrease in entropy; so that if the two compartments are now connected by a heat engine, we seem to obtain a perpetual-motion machine of the second kind.

It is simpler to repel the question posed by the Maxwell demon than to answer it. Nothing is easier than to deny the possibility of such beings or structures. We shall actually find that Maxwell demons in the strictest sense cannot exist in a system in equilibrium, but if we accept this from the beginning, we shall miss an admirable opportunity to learn something about entropy and about possible physical, chemical, and biological systems.

²⁸ *Cybernetics* 71-3.

For a Maxwell demon to act, it must receive information from approaching particles, concerning their velocity and point of impact on the wall. Whether these impulses involve a transfer of energy or not, they must involve a coupling of the demon and the gas. Now, the law of the increase of entropy applies to a completely isolated system, but does not apply to a non-isolated part of such a system. Accordingly, the only entropy which concerns us is that of the system gas-demon, and not that of the gas alone. The gas entropy is merely one term in the total entropy of the larger system. Can we find terms involving the demon as well which contribute to this total entropy?

Most certainly we can. The demon can only act on information received, and this information ... represents a negative entropy. The information must be carried by some physical process, say some form of radiation. It may very well be that this information is carried at a very low energy level, and that the transfer of energy between particle and demon is for a considerable time far less significant than the transfer of information. However, under the quantum mechanics, it is impossible to obtain any information giving the position or the momentum of a particle, much less the two together, without a positive effect on the energy of the particle examined, exceeding a minimum dependent on the frequency of the light used for examination. Thus all coupling is strictly a coupling involving energy; and a system in statistical equilibrium is in equilibrium both in matters concerning entropy and those concerning energy. In the long run, the Maxwell demon is itself subject to a random motion corresponding to the temperature of its environment, and as Leibnitz says of some of his monads, it receives a large number of small impressions, until it falls into 'a certain vertigo', and is incapable of clear perceptions. In fact, it ceases to act as a Maxwell demon.

Nevertheless, there may be a quite appreciable interval of time before the demon is deconditioned, and this time may be so prolonged that we may speak of the active phase of the demon as metastable. There is no reason to suppose that metastable demons do not in fact exist; indeed, it may well be that enzymes are metastable Maxwell demons, decreasing entropy, perhaps not by the separation between fast and slow particles, but by some other equivalent process. We may well regard living organisms, such as Man himself, in this light. Certainly the enzyme and the living organism are alike metastable: the stable state of an enzyme is to be deconditioned, and the stable state of a living organism is to be dead. All catalysts are ultimately poisoned: they change rates of reaction, but not true equilibrium. Nevertheless, catalysts and Man alike have sufficiently definite states of metastability to deserve the recognition of these states as relatively permanent conditions.

A comparative grammar of the Hittite language, revised edition. By EDGAR H. STURTEVANT (and E. ADELAIDE HAHN). (William Dwight Whitney linguistic series.) Vol. 1, pp. xx, 199. New Haven: Yale University Press, 1951.

Reviewed by HOLGER PEDERSEN, *University of Copenhagen*

This book is a revised and completely rewritten version of a work first published in 1933. In spite of the appearance of E. Adelaide Hahn's name on the title page, the book is wholly by Sturtevant; Miss Hahn is to be the author of a projected second volume, which will treat the syntax of Hittite.¹ This first volume contains

¹ [Cf. E. Adelaide Hahn, Lg. 28.422 fn. 19.]

the phonology and morphology. Let it be said at once that the work before us is a very notable achievement: we have here the result of a long and intensive study of the Hittite texts, beginning about 1924 and ending only shortly before the author's death.

No doubt the chief task of a reviewer is to call forth discussion. But in this short review it would be out of place to argue the merits of Sturtevant's Indo-Hittite hypothesis. I will only say that if, in what follows, I sometimes use Sturtevant's terminology (distinguishing, for instance, between Proto-Indo-European and Proto-Indo-Hittite), it is not because I have given up my own views. Aside from such theories, it is evident that in documents as early as those of Hittite, many forms will be found that antedate the oldest IE forms hitherto known; but at the same time, because of the rapid development of the Hittite language, there will be later forms also.

The development of IE *o* before *n*, *m*, *r*, a question on which I cannot agree with Sturtevant, I shall treat elsewhere; here I shall discuss instead a problem of consonantism. Sturtevant is right to give up (57, §80) the theory advanced in his first edition, 'that IH *gw* and *g'w* yielded Hitt. *w*'. This theory was unlikely a priori; for the voiceless, voiced, and aspirated stops normally fall together in Hittite, except that the voiceless stops are written double between vowels. IH *gw* and *g'w* should therefore yield simply *kw* (undoubled); and in fact we find *ku-en-zi* 'he strikes', cognate with Skt. *hanti* 'kills' and Gk. *θείω* 'strike', *φόνος* 'murder'. That is to say, Hittite in this respect agrees completely with the other centum languages, Greek, Latin, Germanic, and Celtic; it cannot teach us anything more about the stops of the proto-language than we have already learned from these branches. So far so good. But Sturtevant has committed a fatal error in writing *kw*, *gw*, *ghw* with a full-size *w* instead of with Brugmann's modest superior *ʷ*. Sturtevant takes his *w* seriously: he tells us that after IE *k* and *ĸ*, *g* and *ĝ*, *gh* and *ĝh* 'there may stand a *w/u* in Hittite and in the IE languages' (56, §78)—in other words, that this *w* sometimes appeared, sometimes failed to appear. On this basis, he has no hesitation in connecting *ku-ir-zi* 'cuts' with Gk. *κείρω* 'cut' (but the suggestion that *κείρω* has lost a labial element by dissimilation in *κρέας*, Skt. *kravis*, is not worthy of Sturtevant). If Sturtevant would convince us that this is anything more than a case of phonetically similar words with approximately similar meanings, he should give us many more instances of the same kind; and it would have been interesting to see at least one instance of the parasitic *w* after *ĸ*. In short, there is no doubt that Brugmann was right in positing three orders of *k*-sounds. Two of these are easy to locate: the *ĸ*-sounds must have been pronounced in the *i* region of the tongue, the *kʷ*-sounds in the *u* region. As for the third order, since it cannot have been farther front than *ĸ* and *kʷ*, it must have been farther back; hence it must have consisted of true velars, and we are justified in speaking of a *q*-series. (But Brugmann ought not to have called the middle series 'labiovelars'; they were not velars at all.) That this system was liable to simplification is easily seen. The back-vowel timbre of the *kʷ*-sounds could develop into a full consonantal *w*, whereupon the velar would be free to become a normal *k* and the *ĸ* to give up its front-vowel timbre. By a different but equally probable development, the *ĸ*-sounds could develop into *t*-sounds and *š*-sounds, leaving the *kʷ*-sounds free to give up their back-vowel timbre and

develop, together with the velars, into ordinary *k*-sounds. Thus it is that *k̂* and *q* have yielded the same reflex in Greek and Latin, whereas in Sanskrit it was *q* and *k̂* that fell together. In Irish all three orders of voiceless and aspirated *k*-sounds became one (there is no need to describe here the more complicated development in Brythonic); but the voiced *q̂* retained its labial element (cf. Ir. *ben* 'woman' etc.). Curiously enough it is the Albanian language which to this day has kept alive traces of the difference between *q* and *k̂*; cf. my remarks in KZ 36.305, 340, and A. Mayer in KZ 70.87 and fn. 1. It is clear now that Albanian points to an original velar quality for *q*, contrasting with more forward position of *k̂*.

There are a good many etymologies in Sturtevant's *Grammar* that struck me as highly dubious. I shall not enumerate them; but I cannot refrain from remarking that when Sturtevant sees the root **sed-* in Hitt. *ḫa-aš-du-e-ir* 'boughs', Gk. *ᾠος* (65, §87), he falls below his usual high standard.

Sturtevant posits four laryngeals, which he writes ' , *h*, *x*, *γ*. The development and effect of these sounds are described in §§74-77a (47-55). We learn that IH *x* becomes Hitt. *ḫḫ* between vowels; that IH *e* next to an *x* was pronounced *a*; and that *γ* was never written double in Hittite (cf. *me-ḫur* 'time'; Skt. *māli* 'measures', Gk. *μηρίς*, Lat. *mērior*, Goth. *mēl* 'time'). Sturtevant tells us (52, §76) that *h* 'is lost in Hittite and the other Anatolian languages, as in Indo-European, but in both it has induced a change of *e* to *a*', and (54, §77) that 'IH ' has no effect upon vowel quality either in Hittite or in the Indo-European languages. It lengthens a preceding vowel upon its loss before a consonant in pre-IE.' The arguments that have led Sturtevant to assume the existence of the two laryngeals ' and *h* in some remote period of pre-Indo-Hittite are undeniably ingenious; but they go far beyond the methods of comparative linguistics in the proper sense. Leaving speculation aside, we see no trace of these two laryngeals, but only certain features of ablaut.

The other two laryngeals, on the other hand, namely *x* and *γ*, are guaranteed by the results of strictly linguistic comparison. In this connection, Sturtevant's interpretation of Skt. *pibati* 'drinks' (51, §75 end) is highly interesting: he derives it, along with Lat. *bibit* and OIr. *ib-*, from IE *pibēti* < IH *py-by-ēty*.

Seventeen pages of the book (149-65) are given over to verbal paradigms. I wish that Sturtevant had followed these with a complete list of Hittite verbs (including the denominative), with citations of all their occurrences in the *Grammar*. If he had done that, he need not have cited so many different forms for each verb in his Index. As for this Index itself, it seems to me most inconveniently arranged. There are three columns per page: the first shows the syllabic transliteration of the Hittite words, the second shows the suggested phonetic interpretation, the third refers to the sections in which the cited form appears. Without the use of a ruler laid across the page to lead the eye from column to column, it is often scarcely possible to use the Index at all. It would obviously have been a more useful plan to give the syllabic transliteration and the phonetic respelling together (the latter in brackets), and immediately thereafter the section numbers. In this way, incidentally, two columns on a single page could have contained what is now spread wastefully over two pages.

Errors in the Index are especially troublesome, and unfortunately they are not

rare. In a whole series of references (under *da-at-ti*, *a-a-ri*, *a-a-an*, *a-a-an-za*, *a-ap-pa*, *u-uh-ḫi*, *pa-a-i*, *pa-it-ti*, *pa-iš-ta*, *ša-ra-a*, *u-*, *ša-ak-ti*, *ša-ak-ta*, *ša-ra-at-ti*, *ša-ra-at-ta*), we find '76a' instead of '76'. Under *e-ku-(i-e)-ir* the reference '237' is wrong; perhaps it should be '77'. Similar inaccuracies—to say nothing of frequent misspellings and distracting typographical inconsistencies—appear throughout the text as well. Thus, the reference in footnote 101 (54) to 'Pokomy, Idg. et. Wörterb. 23' should be instead to page 34—and the name, of course, should be 'Pokorny'. A really bad misprint is 'Lyc. uga-' for *huga-* (47, §74).

I was occasionally misled by Sturtevant's practice of omitting the comma between coordinate forms, e.g. 'the enclitic reflexive *-az -za*' (28, §53), *da-a-an ta-a-an* (61, §83), *lu-ú-lu lu-lu-ú* (68, §90). The last example is especially irritating, since it occurs in a section on reduplication, and might easily be mistaken for a single word like the *har-ši-har-ši* cited a few lines earlier.

In spite of these shortcomings, I am convinced that the publication of Sturtevant's Hittite grammar in this revised edition is one of the happiest and most significant events in the recent history of linguistic science.²

Die englische Sprache: Ihre geschichtliche Entwicklung. By KARL BRUNNER. (Sammlung kurzer Grammatiken germanischer Dialekte: B. Ergänzungsreihe, Nr. 6.) 2 vols., pp. xix, 352, 424. Halle (Saale): Max Niemeyer Verlag, 1950, 1951.

Reviewed by HERBERT PENZL, *University of Michigan*

Karl Brunner, Professor of English in the University of Innsbruck, has in recent years published revisions of Eduard Sievers' venerable *Angelsächsische Grammatik* and of his *Abriss der angelsächsischen Grammatik*, as well as an *Abriss der mittenglischen Grammatik* of his own (2d ed. 1948). Here he has presented us with a bulky but useful manual on the history of the English language. The first volume (*Allgemeines; Lautgeschichte*) contains chapters on the origin of English (1–38), which treat the settlement history of England; on related languages (41–80), discussing the phonological and morphological features of Indo-European, West Germanic, and Anglo-Frisian; on the written tradition and 'areal subdivision' of English (81–116); and on foreign influences (117–193), dealing with the entire external and internal history of North Germanic, Anglo-Norman, and other foreign languages on English soil. The two final chapters take up the development of English vowels in stressed (194–284) and unstressed syllables (284–302), and that of English consonants (303–51). The second volume contains two major chapters: *Die Flexionsformen und ihre Verwendung* (1–344), *Die englische Sprache ausserhalb Europas* (345–83). The treatment of morphology includes the following parts of speech: substantives (2–49); adjectives and adverbs (49–80); numerals (80–93); pronouns (93–157); verbs (157–344). The work ends with a valuable *Sachregister* (384–9) and *Wortregister* (390–421), which give page instead of section references to both volumes.

² It is perhaps not out of place to mention here another publication, issued just as I was finishing this review: Emmanuel Laroche, *Recueil d'onomastique hittite* (Paris, 1952). At least one chapter in this book, entitled 'Suffixes de dérivation', has direct bearing on the grammar; but even the lexical materials which are its chief contents may prove to be of use to the grammarian.

In dealing with morphology, Brunner first gives the OE forms in a brief comparative Germanic (and Indo-European) setting—a welcome procedure, rarely found outside of Kieckers' *Altenglische Grammatik*. He traces their phonetic development in ME dialects, gives bibliographical references to special studies, and discusses in detail their syntactic occurrences in Modern English, and their meanings from the viewpoint of a speaker of German. The sources for the MnE examples are usually indicated. Brunner's sound and highly readable account of the historical development of all major morphological units makes the chapter on morphology very useful for purposes of reference.

As for his chapter on the English language outside Europe, it must be assumed that some statements about American English would look different if the author had written them after his recent trip to the United States; in Austria he had only secondary sources and no informants at his disposal. The 'broad a' in *half*, *past*, *dance*, current in parts of New England and of the South, is certainly not, as Brunner states (2.363), just an imitation of Standard British English, nor is the long vowel in *dance* [dæ:ns] an 'Art Kompromiss mit der südeagl. Lautung' (1.265).

Some scholars in this country will probably consider it a serious shortcoming in Brunner's work that he makes no attempt to deal with historical morphology, nor even with historical phonology, from the structural point of view. Brunner rarely uses the term 'Phonem'; only once among four occurrences do we find it without quotation marks. Instead, Brunner's chapters on phonology exemplify the point of view found in Karl Luick's great *Historische Grammatik der englischen Sprache*. Luick's description of the Modern English vowel shifts introduced the vowel patterns as frames of reference. Like Luick, Brunner (1.56 f.) adopts Paul Kretschmer's somewhat teleological interpretation of the Germanic consonant shift, which seems to reveal the influence of the Prague phonologists. All this may not add up to an impressive record of familiarity with modern phonemic theory, even for the spring of 1945 (when the manuscript of the phonological part was finished), but it would be unrealistic of us to expect more of a break with a distinguished tradition. Besides, it must be granted that the structural restatement of minor chapters in the history of the English language entails overwhelming difficulties, since for a comprehensive treatment the necessary preliminary studies are lacking, and the application of rigorous structural techniques to historical material is, even in this country, still in its infancy.

Any real advance in historical phonology will necessitate a careful revaluation of methods as part of a consistently structural approach. The question will have to be asked and answered for each body of material: what type of evidence can be relied upon for what type of phonetic change? A discussion of such basic questions would seem to be important in a work intended as a textbook for university students, since an understanding of general linguistic principles is certainly no less essential as a pedagogical objective than familiarity with specific historical facts of linguistic development. Methodological clarification, moreover, is not only a pedagogical but especially a scientific desideratum. Brunner lists (112-6) the types of evidence available for phonemic change: occasional spellings ('gelegentliche Schreibungen'), orthoepic statements, rimes, loanwords; in dis-

cusssing specific phonemic changes, he usually quotes the pertinent evidence, at least briefly. But he does not characterize or discuss the methodology involved.

Several writers have advocated caution in the use of naive or occasional spellings as evidence for a general phonemic or phonetic change. Brunner, on the whole, does not place much emphasis on them; but apparently he does not hesitate to take OE *ui* spellings in place of *y* at their face value as evidence for a pronunciation [uy] for [y] (1.203).

The correlation between an orthographic change and a phonetic change is not always as obvious as most scholars seem to assume. The gradual replacement of OE *æ* by ME *a* is an orthographic change that has usually been interpreted as indicating a phonetic change of [æ] to [a] rather than a phonemic change, i.e. a coalescence of the OE phonemes /æ/ and /a/ into ME /a/. It is this obsession with the specific phonetic values of written symbols that in the past has consistently obscured the more important phonemic facts of patterning and contrast. The ME symbol *a* does not reveal to us any particular sound quality; it may cover a multitude of allophones. The modern value [æ] may indeed have been in existence since OE times. The assumed sequence of sound-changes: OE [æ] to ME [a] (Brunner 1.213) to Early MnE [æ] (Brunner 1.255), has rightly been viewed with suspicion by many scholars. Only orthoepic evidence, which is often highly dubious, is available for the alleged MnE phonetic change.

Brunner's book was published in East Germany under the most difficult circumstances. This fact is reflected in the poor quality of the paper and in a comparatively high number of misprints (most of them corrected under *Nachträge* and *Berichtigungen*: 1.352, 2.422-4). His style and his manner of presentation are admirably clear and straightforward. The scope of his work, the thoroughness and competence of his treatment make it a valuable reference book for all students of the history of English.

Esquisse de la langue hongroise. By AURÉLIEN SAUVAGEOT. (Les langues et leurs structures, No. 3.) Pp. 336. Paris: Librairie C. Klincksieck, 1951.

Reviewed by THOMAS A. SEBEOK, *Indiana University*

'Phoneme' was the rallying cry of the new linguistics through the thirties, climaxed by the publication in 1939 of Trubetzkoy's classic yet now curiously obsolete magnum opus. If the key term was 'phoneme' then, it is now, in the early fifties, assuredly 'structure'. But if 'the phonemic principle is accepted by all modern schools of linguistics', as has been hopefully claimed,¹ the structural principle remains as yet largely unassimilated. This is true not only in linguistics, but in modern scientific thinking in general. The British physicist Lancelot Law Whyte, discussing the way in which similar ideas emerge in many different realms of thought,² continues: 'It seems that a major movement of this kind is now affecting many branches of science: the decline of what may be called *atomism*, or atomistic thought in general, and the emergence of theories based on the conception of *patterns* . . . It is being realized that the structure of natural (and

¹ Lotz, JASA 22.715 (1950).

² Harper's Magazine, February 1950, 25-6.

human) processes may not be reducible to the haphazard interactions of unchanging units, but may involve some law of developing pattern. This tendency is unmistakable in many fields though we do not yet know exactly what is meant by a "pattern." Yet it appears that the problem of the scientific philosophy of structure may be rapidly nearing the threshold of a solution. Witness such attempts as Ludwig von Bertalanffy's *Outline of General System Theory*,³ and, perhaps more immediately relevant to linguistic science, the burgeoning of communication theory,⁴ as well as the elaboration of structural ideas in the humanities in general.⁵

In linguistics proper, 'structure' is commonplace in the titles of current books, collections, and series: Harris calls his work *Methods in structural linguistics*; the Hjelmslev Festschrift is called *Recherches structurales*; Trager and Smith have named their recent monograph *An outline of English structure*; the Linguistic Society publishes in its series of Language Monographs a subseries called *Structural Sketches*; and the book under review constitutes the third in a series of volumes bearing the overall title *Les langues et leurs structures*. (The first of these, by the same author, dealt with Finnish;⁶ the second, by Mirambel, to deal with Modern Greek, is announced for later publication.) Under the circumstances, it seems not amiss to inquire whether we are talking about the same thing. Does the American series have the same object in view as the corresponding French series? Martinet's acid reflections *About Structural Sketches*⁷ would incline one to anticipate considerable differences in approach.

If a student attempts to discover what the professional linguist means by 'structure', he soon finds a range of usage which varies widely in rigor. On the one hand there are attempts at precise definition of structure, for example in some of Hjelmslev's writings; at the other extreme there are mere expressions of a vague emotional identification with what is considered a prestige group. If Sauvageot's presentations of Finnish and Hungarian are representative samples of this new French series, then the term 'structure' is used here with something less than rigor, and certainly not as it is used in American combinatorial linguistics, or in the writings of the Prague school, or in glossematics.

This book of Sauvageot's is not a scientific grammar in which forms are identified and accounted for in their systematic relationships, but rather an artistically selective assemblage of traits in the Hungarian language that appeal to the author as in some way essential (7). His basis for deciding what is relevant is unfortunately suspect on several grounds. In the first place, Sauvageot suffers from the bias that Hungarian and Finnish have the same morphology;⁸ thus, the major plan, as well as numerous details of the design of the Hungarian sketch, except for an additional section on phrases (*Les groupes de mots*, 93-155), is

³ The British journal for the philosophy of science 1.1-32 (1950).

⁴ See the excellent synthesis by Jurgen Ruesch and Gregory Bateson, *Communication* (New York, 1951), esp. ch. 10 and Tables A-C (259-61).

⁵ Cf. the reviewer's discussion, *The structure and content of Cheremis charms*, *Anthropos* (forthcoming).

⁶ Reviewed in *Lg.* 23.173-8 (1947).

⁷ *Word* 5.13-35 (1949).

⁸ *Actes du sixième congrès international des linguistes* 454 (Paris, 1949).

identical with that of the Finnish (see Lg. 23.174 f.), resulting from a deliberate attempt (7) to emphasize resemblances. All this would lead one to conclude that the author's acquaintance with these languages is rather superficial; as Lotz has remarked,⁹ 'the idea of similarity is extremely vague in Sauvageot's usage—many of the facts he presented have to be analyzed more intensely to achieve a complete description . . .'

In the second place, Sauvageot seems unable to escape the bias of his meta-language. This imparts, to be sure, a certain Gallic charm to the book, as can be illustrated by portions of a paragraph (13) describing the acoustic impression supposedly made by Hungarian: 'Pronounced in a grave manner, articulated rather far behind the teeth, the mouth being either boldly open or outright rounded, Hungarian often sounds a bit sepulchral . . . it gives the impression of a sort of psalmody . . . but always virile'. It also explains why a negative statement to the effect that the Hungarian noun has no gender can be immediately followed by a two-page discussion of 'le genre' (55-6).

The analysis is thus neither comprehensive nor deep; in consequence and not unexpectedly, it is studded with inaccuracies and half-truths. To take as an example only a single typical paragraph dealing with a relatively simple matter, viz. the inventory of Hungarian fricatives, it is hard to find in it a single statement that is not inaccurate or confused (or at best not obsolete): we read that *f* is bilabial; that *j* is sustained longer than in French; that *h* is less aspirated than in German and that between vowels it resembles a *v*; and that the sound-type [ç] is 'gutturale mouillée qui ne fait son apparition qu'à la fin du mot', though it is neither 'guttural' nor 'palatalized', and is not limited to word-final position. True to the title of the book, this is certainly sketchy; but is it Hungarian?

The book, in fact, adds nothing to our knowledge of Hungarian, nor does it sharpen our tools of analysis. It is not even useful as a pedagogic device. Unlike its Finnish counterpart, it seems to have been composed too mechanically to achieve an effect of coherence. It comes to little more, in fact, than a potpourri of impressions dished up for the benefit of the unsophisticated French reader.

If this is structural analysis, make mine old-fashioned!

* Ibid. 465.

Notes on Gurage grammar. By H. J. POLOTSKY. (Oriental notes and studies, No. 2.) Pp. 58. Jerusalem: Israel Oriental Society, 1951.

Reviewed by WOLF LESLAW, *Brandeis University*

The study of the Ethiopic languages of Gurage, spoken to the southwest of Addis Ababa, is à l'ordre du jour. There are three main groups in Gurage: (1) Eastern Gurage, including Selti, Wolane, Ulbarag, and Zway; (2) Western Gurage, including Chaha, Ennemor, Eža, Endegeñ, Muher, Gogot, and Masqan; and (3) Northern Gurage, with Aymellel as the only representative. The first study of Aymellel was published by Johannes Mayer,¹ later elaborated by F. Praetorius.² A collection of words and some grammatical notes in the dialect

¹ *Kurze Wörtersammlung in Englisch, Deutsch, Amharisch, Gallanisch, Guragesch*, ed. L. Krapf (Basel, 1878).

² *Über den Dialekt von Guragué, Die amharische Sprache* 507-23 (Halle, 1879).

of Chaha were published by G. Chiarini.³ A summary description of this dialect was given by C. Mondon-Vidailhet,⁴ and a small vocabulary of Chaha, Gogot, and Ulbarag, collected by the same author, was edited by E. Weininger.⁵ To Marcel Cohen we owe a brief but excellent description of Eža (wrongly considered by him to be Chaha), Muher, Aymellel, and Wolane, a vocabulary of these four dialects, and a few texts.⁶ Some particular problems of Eža (again mistaken for Chaha) were the subject of an article by H. J. Polotsky.⁷

In view of the great variety of the Gurage dialects and their important place in Ethiopic dialectology, I decided to investigate as many of them as possible during my trips to Ethiopia in 1946-7 and in 1950.⁸ I studied the general structure of nearly all the dialects, and collected a vocabulary of over 1000 roots for eleven of them. For Chaha, the language of what is perhaps the most important Gurage tribe, I published a collection of texts describing various phases of Gurage life, folktales, proverbs and some songs; see my *Ethiopic documents: Gurage* (VFPA No. 14, 1950). This publication also contains an outline grammar and a vocabulary of Chaha. The grammar, limited to twenty pages, proved to be a difficult task. Chaha is one of the most complicated languages of the Semitic group in Ethiopia. Since the purpose of my outline was 'to make understandable the forms that occur in the texts', there was no special reason to enumerate all cases of what I call umlaut and Polotsky calls 'i-epenthesis' (23), or all the ways in which the gerundive is expressed (41 ff.). In the booklet under review, Polotsky in a masterly way clarifies various features of the language and corrects a number of errors in my outline. His observations on the phonemic problems of Chaha vowels, on labialization, on the object and mediate suffixes, and the like, all based on my texts, are judicious and penetrating, and will prove most useful for future studies of Chaha and of Gurage in general.

Some remarks on Polotsky's *Notes* follow. References are to pages: the numbers in parentheses refer to sections of my outline.

15 (2.1): To the cases in which *a* and *i* are in contrast, add the imperative masc. *sədab* 'insult', fem. *sədib*.

20 (3.1): I stated that Chaha has no gemination, but that traces of primitive gemination remain in the devoicing of primitive voiced geminates, with aspiration of dentals, labials, and velars. Careful examination of these sounds during my second stay in Ethiopia convinced me that my statement about the aspiration was wrong. These sounds are only occasionally aspirated, and the aspiration is certainly not phonemic.

22 (4.2): The root for 'accuse' is indeed *k'äsäsäm*, not *käsäsäm*. I recorded this root with *k'* also in Eža and Ennemor; the impersonal is *k'äšäšim*, not *k^wäšäšim*.

23 (5): Since Polotsky has raised the question of *i*-epenthesis, I should like to add that there are many more cases than those that he mentions. Thus, with a

³ Note grammaticali e vocaboli della lingua ciahā (guraghè), in Antonio Cecchi, *Da Zeila alle frontiere del Caffa* 3.469-84 (Rome, 1887).

⁴ *La langue harari et les dialectes éthiopiens de Gouraghè* 77-119 (Paris, 1902).

⁵ *Études sur le gouragiè* (Wien, 1913).

⁶ *Études d'éthiopien méridional* 55-241 (Paris, 1931).

⁷ *Études de grammaire gouragué*, BSL 39.137-75 (1938).

⁸ A year of research in Ethiopia, Word 4.212-25 (1948); Report on a second trip to Ethiopia, Word 8.92-9 (1952).

last labial **omi* becomes *em*⁹, as in *tətom* 'you fast' (masc.) : fem. *tətem*⁹. Note also **ūri* > *e*, and **ari* > *e*, *ā*, as in *təwār* 'you (masc.) spend the day' : fem. *təwe*; *čar* 'load' (masc.) : fem. *če*. I did not record prepalatalization or *i*-epenthesis in type-B verbs with *n* as last consonant (of verbs with an original last radical *y*), e.g. *šänäm* 'feed' : imper. masc. *šän*, fem. also *šän*. Polotsky is right in stating that a pattern *qətlä* does not exist: the form *səḏäbä* represents *qətale* with *i*-epenthesis. For that matter, *qətaläyä* (as given in my §9.3) does not exist either: the form *nəmağä* 'love' represents *qətale* (**nəmade*) with prepalatalization of *d* to *ğ*.

23 (5): The imperfect sg. 2nd fem. of verbs in *-a* with a labial as last consonant is indeed *təsemä*, *təsebä*, *tətefä*, as against the masc. *təsäma*, *təsäba*, *təläfa*, etc.

24 (5): It is true that for the jussive sg. 2nd fem. one finds *tənṣäb* (Polotsky's *tənṣäb*); but I also recorded *atənṣ'äb* beside *atənṣäb* for the negative jussive, and *nəṣäb* beside *nək'äb* for the imperative.

34 (28.1): For the copula 2nd fem. sg. after a consonant I recorded *nəx* in *mäman nəx* 'how art thou?' (not *ənəx*, as Polotsky rightly points out), but also *ənəx* as in *ənəx* 'thou art tall'.

34 (32.2): Concerning the three types of the jussive it is true that *A₃ yäqəll* was originally a phonetically conditioned variant of *A₁ yäqəll*; but as I have tried to show elsewhere,⁹ this is no longer the case. *A₃* is now a jussive type independent of *A₁*.

35 (34.2b): The origin of *sarä* 'be beautiful', with object suffix pronouns 'be contented, pleased',¹⁰ is Geez *šāhalä* 'be gracious'. Its imperfect *yəšər* (with prepalatalized *š*) can be explained by **yəsəhal* > **yəsəyal* > *yəšər*, like the imperfect *yəxər* < **yəkəyal* < **yəkəhal* given by Polotsky. One might ask, however whether the perfect *sarä* does not represent the stage of Geez when *š* became *s*, whereas the imperfect preserves the archaic sound.

35 (38.3): It is true that 'the nature of the root constitution' is apparent in the causative morpheme *at-* (not *a-*) in verbs with initial *a-*; but in the regular verb, my statement that 'nothing in the nature of the root constitution or in the meaning of the basic stem can determine whether the prefix is *a-* or *at-*' holds true. The causative morpheme *at-* seems to be added, as a rule, to the base of type B, 'irrespective of the type of the underlying simple stem' (36); but I also recorded *at-bäsäräm* (from the base of type A, not *at-besäräm*), *at-säfäräm* (type A, not *at-sefäräm*), *at-qänäsäm* (type A, not *at-q'anäsäm*).

42 (46.6): To the tenses of verbs with an 'adverbial' function, such as *täzä-päräm* 'return', *käbäm* 'repeat', *abätäräm* 'do first', add the gerund form (45) of *käbäm* in the sentence *käb-tä yəčän-te* 'he will come again'.

45 (46.6): The particular gerund form does indeed look like the imperative fem. sg. in all verbs, not only 'in most cases'. For the added element read *tä* instead of *ät*. For the 1st person I also recorded *-tuḵ* beside *täwḵ*. Unlike the Chaha Catechism¹¹ I found *tä* for all persons: sg. 3m. *tä*, fem. *täč*; 2m. *täka*, f. *täx*; 1st *tuḵ* and *täwḵ*; pl. 3m. *täbo*, fem. *täma*; 2m. *täku*, fem. *täkma*; 1st *tänä*. The

⁹ Le type läbsä en gouragué, *Rassegna di studi etiopici* 10.85-98 (1951).

¹⁰ For the relation between these two meanings, see JAOS 71.227 (1951).

¹¹ *Ačər temherlä krəstyan bəğurage bāčāha q'anq'a* (Dire-Daoua, 1933).

complex of this particular gerund form + *tä* takes the object suffix pronouns: *qarəčə-te* 'he despising me, when he despises me', *qarəče-tä-nax* 'he despising thee (fem)'.

53 n. 63: The paradigm of the impersonal imperfect of *dänägäm* has actually been heard from an informant. For the imperfect I also recorded 2nd masc. sg. *yədärgukä*, 1st person *yədärgun*, pl. 2nd masc. *yədärguku*, fem. *yədärgukma*, 1st *yədärgundä*.

55 n. 80: I fail to understand the problem of the vowel of the 2nd radical of verbs with root-final laryngeal. Polotsky speaks here of 'Primitive Western Gurage', and mentions the second *ä* of Muher *fäkkä-m^w-em* and Aymellel *fäkkä-mu-n*. Of all the Western Gurage languages, however, only Muher and Gogot have *ä* in the 3rd masc. pl., a form in which the old *-u* was replaced by *-m^w*: Gogot *märrä-mum* (for Muher I recorded *sämmä-m^wəm^w*). The other languages of the Western group replaced the old *-u* by *-o*, without *-m*: Chaha *gäpo-*, Masqan *märro-*. I wonder, therefore, whether one is entitled to speak of a form of 'Primitive Western Gurage'.

Polotsky's *Notes* make profitable reading, and help to elucidate a number of obscure problems in Gurage dialectology.

NOTES

JAKOB JUD, Professor of Romance Linguistics in the University of Zurich, and an Honorary Member of the Linguistic Society of America since 1937, died in Zurich on 15 June 1952, at the age of seventy.

Jud was trained in Romance philology at the University of Zurich, where he took his doctorate in 1907, and at Paris, where he studied linguistic geography under Jules Gilliéron. From 1907 to 1915, he was Privatdozent at Zurich; from 1915 onward, he was professor and head of the Romance Seminar. In the summer of 1931, he was Visiting Professor in the Linguistic Institute, held that year in New York; he gave a course on problems in the preparation of linguistic atlases, and helped to train the staff of the *Linguistic Atlas of New England*.

The generation of Gilliéron's disciples to which Jud belonged was a brilliant one, including such men as Karl Jaberg, Matteo Giulio Bàrtoli, and Antòni Griera. All of these carried away from Gilliéron's teaching an enthusiasm for linguistic geography which led them to apply his methods to hitherto unexplored aspects of Romance dialectology. Jud's chosen field was the Rhaeto-Romance and North Italian dialect area, as evidenced already in his first major publication (*Zur Geschichte der bündnerromanischen Kirchensprache*, 1919). His masterwork is, of course, the immense *Sprach- und Sachatlas Italiens und der Südschweiz* (AIS) (1705 maps in 8 volumes, 1928-40, with three subsidiary volumes—introduction, illustrations, and index, the last still to appear). In addition to this magnum opus, he was the author of numerous articles on Gallo- and Ibero-Romance dialectology and word history, and was editor of the periodical *Vox romanica* from 1936 on. At one time he also planned a Romance etymological dictionary, along semantic lines; unfortunately this project never came to fruition, and remains in the shape of extensive files of slips.

But Jud was considerably more than a mere imitator of Gilliéron; in his work, he carried on and developed Gilliéron's approach, both refining it and reintegrating it with the previous findings of more traditional Romance philology. The AIS represented a marked improvement over the French atlas in a number of respects, such as choice of informants, questionnaire, and presentation of the material. Jud's work in this field is of especial importance to American linguistics, since it was the point of departure for the development of the American linguistic atlas. In his own research and writing, Jud combined the techniques of linguistic geography with the historico-philological approach (as in: *Zur Geschichte der romanischen Reliktwörter in den Alpenmundarten der deutschen Schweiz*, VR 8.34-109 [1945/6]) and with comparative reconstruction (as in: *Altfrz. estuet; bündnerrom. stuwer, stuwair*, VR 10.29-56 [1948]). In this respect, Jud contributed more, perhaps, than any other single scholar of his generation to making linguistic geography an integral part of the technique of Romance linguistics, rather than a discipline opposed to or even superseding other approaches, as Gilliéron and some of his followers had tended to conceive it.

Jud's death is a severe blow to Romance linguistics; but he has left with us, not only the actual corpus of his publications, but the memory of his kindly, cheerful,

invigorating personality, and above all the lasting contribution he has made to the methods of our science.

ROBERT A. HALL JR., *Cornell University*

THE SEVENTH INTERNATIONAL CONGRESS OF LINGUISTS was held in London, 1-6 September 1952, in the stately Senate House of London University. The Linguistic Society of America was represented by three official delegates: Bernard Bloch, Hans Kurath, and Joshua Whatmough. (Two of these were present also as members of the Comité International Permanent de Linguistes: Whatmough as successor to Leonard Bloomfield, and Bloch as proxy for W. Freeman Twaddell.) In addition, at least twelve other members of the Society made the transatlantic journey to attend the Congress: Yehoshua Bar-Hillel, Madison S. Beeler, William S. Cornyn, Pierre C. Delattre, Charles C. Fries, Eric P. Hamp, W. Cabell Greet, William N. Locke, William F. Mackey, André Martinet, Ernst Pulgram, and Harry V. Velten. In all, there were approximately four hundred registered delegates from forty different countries (none from east of the Iron Curtain).

The program followed, in general, traditional lines, with most of the sessions devoted to the comparative and historical study of well known language families, or here and there to more detailed concern with particular languages. Some time, however, was given to considering much more general topics, in particular the nature of meaning, the requirements of a conceptual dictionary, statistical linguistics, and synthetic speech—the last, as presented by Delattre, in many ways the high point of the entire Congress. Alf Sommerfelt, speaking for the Comité International Permanent de Linguistes, gave an informal report on the activities of the Comité, its role in the organization of future Congresses, and a plan to reconstitute the Comité so as to secure wider representation of different countries. In summarizing the activities of the present Congress, Sommerfelt called attention to the lively and pervading interest shown by scholars from all parts of Europe in American structuralism, and expressed the hope that international cooperation among linguists would continue to grow.

It may perhaps be regretted that the scope of the topics discussed was not broader. The method of presenting papers prepared in advance, followed here as in nearly all international congresses, is certainly to be condemned, since it prevents continuity and coherence and thus effectively blocks real discussion. As far as was possible within the limits imposed by this method, however, the organizers of this Congress tried to encourage a free exchange of opinions.

In spite of serious difficulties created by the last war, the London Congress succeeded in preserving the high standard of previous Congresses.

A BIBLIOGRAPHY OF STATISTICAL LINGUISTICS is now being finally revised at Harvard University, under the direction of Joshua Whatmough, for eventual publication by the Comité International Permanent de Linguistes. In order that this work may be as complete as possible, members of the Linguistic Society who have noted recently published material within this field (since about the

end of 1950) are invited to send a brief listing to be inserted in its proper place in the *Bibliography*. Such listing should give the author, title, place, and date; it should be sent to Professor Joshua Whatmough, 54 Dudley Hall, Harvard University, Cambridge 38, Mass.

ANDRÉ MARTINET calls attention to an erroneous illustration in his article, Celtic lenition and Western Romance consonants, *Lg.* 28.192-217 (1952). On page 193, lines 24-6 should be deleted and replaced by the following:

the initial consonant of a feminine noun in the nominative is lenited after (say) *cach* 'every'—which means that if the noun normally begins with *t-*, this *t-* will appear as *th-*, thus *túath* 'people', but *cach thúath* 'every people'. In Old Irish, the language

On the same page, line 34 should be changed to read as follows: the feminine form of the nominative of most pronominals ended in a vowel, namely *-ā*,

THE LINGUISTIC INSTITUTE FOR 1953 will again be held in Bloomington, under the joint auspices of the Linguistic Society and Indiana University. The summer meeting of the Linguistic Society will take place on Friday and Saturday, 7-8 August. Further information may be secured from the Assistant Director of the Institute, Professor Thomas A. Sebeok.

The resident faculty of the Institute will include Householder, Sebeok, Velten, Voegelin, and Whitehall; among the visiting members will be Bloch, Emeneau, Hall, Hoenigswald, Hoijer, Joos, Lotz, Osgood, Uhlenbeck, and Uldall—the last two from the Netherlands and from Denmark respectively. Courses of instruction will cover a wide range of Indo-European, Finno-Ugric, Asian, and American Indian languages; the theory and practice of linguistic description, both phonetic and grammatical; and the relation of linguistics to a number of other disciplines. There will be, in addition, faculty seminars devoted to psycholinguistics, to ethno linguistics, to the theory of translation, and to the philosophy of language. Public lectures by Chao, Fries, Harris, Leopold, Leslau, Malkiel, Poppe, and Whatmough will supplement the regular program.

PUBLICATIONS RECEIVED

This listing acknowledges the receipt of recent works that appear to bear on the scientific study of language. No book can be returned to the publisher, nor can the Editor promise that every book received will be reviewed in the journal. Reviews are published as circumstances permit, and copies are sent to the publishers of the works reviewed.

- Abweichende spät- und vulgärlateinische Perfektbildungen. By Frank G. Banta. (Diss. Bern.) Pp. xiii, 124. Freiburg in der Schweiz: Paulusdruckerei, 1952.
- Acta linguistica 6.57-116 (1950-51).
- Africa und Übersee 36.93-192 (1952).
- Aktuelle Sprachwissenschaft: Zeitgeschehen und Zeitgeist im Spiegel der Sprache. By A. Debrunner. (Berner Rektoratsreden 1951.) Pp. 28. Bern: Verlag Paul Haupt, 1952.
- Os animais na linguagem portuguesa. By Delmira Maças. (Publicações do Centro de Estudos Filológicos, No. 2.) Pp. 431. Lisboa: Centro de Estudos Filológicos, 1951.
- Annales Universitatis Saraviensis (Philosophie—Lettres) 1.1-223 (1952). Saarbrücken: Université de la Sarre.
- Annals of the Bhandarkar Oriental Research Institute 32.1-343, i-xxiv (1951).
- Gli antichi Italici, 2d ed. By Giacomo Devoto. (Collana storica, Vol. 29.) Pp. 356. Firenze: Vallecchi Editore, 1951.
- Archiv orientální 19.2-320 (1951) = Les actes des journées scientifiques d'orientalisme 1949.
- Beiträge zur Etruskologie. By Friedrich Slotty. (Bibliothek der allgemeinen Sprachwissenschaft; 3. Reihe, Darstellungen und Untersuchungen aus einzelnen Sprachen.) Vol. 1: Silbepunktierung und Silbenbildung im Altetruskischen, pp. xvi, 207. Heidelberg: Carl Winter, Universitätsverlag, 1952.
- Berceo 7.1-368 (1952).
- Bibliografia de António da Fonseca Soares (Frei António das Chagas). By Maria de Lourdes Belchior Pontes. (Publicações do Centro de Estudos Filológicos, No. 3.) Pp. 127. Lisboa: Centro de Estudos Filológicos, 1950.
- Biblos 27.1-649 (1951).
- Boletim de filologia 12.225-419 (1951).
- Boletim geográfico 8.1387-498 (1951).
- Boletín de filología 6.1-358 (1950-51). Santiago de Chile, 1952.
- Boletín del Instituto Marco Fidel Suarez 1.117-303 (1951).
- Bulletin analytique: Philosophie 6.171-356 (1952).
- Bulletin of the School of Oriental and African Studies 14.1-412 (1952). University of London.
- Celtica 2.1-216 (1952). Dublin Institute for Advanced Studies.
- The classical weekly 46.1-48 (1952).
- Commentationes fenno-ugricae in honorem Y. H. Toivonen. (Mémoires de la Société finno-ougrienne, No. 98). Pp. iv, 391. Helsinki: Suomalais-Ugrilainen Seura, 1950.
- Cultura neolatina 11.197-318 (1951); 12.1-78 (1952).
- Dativstudien: Dativus sympatheticus und dativus comparationis in der norrönen Sprache. By Vemund Skard. (Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo; 2. Hist.-filos. Klasse, 1951, No. 2.) Pp. [viii], 139. Oslo: Jacob Dybwad, 1951.
- Dictionnaire ngbandi (Ubangi—Congo belge): Français-ngbandi, ngbandi-français. By P. Benjamin Lekens. (Annales du Musée du Congo belge: Série in-8°, Sciences de l'Homme; Linguistique, Vol. 1.) Pp. xii, 348. Anvers: Éditions de Sikkel (publié sous les auspices de la Commission de Linguistique Africaine), 1952.
- Dois problemas da língua portuguesa: O infinito pessoal e o pronome *se*. By Theodoro Henrique Maurer Jr. (Universidade de São Paulo: Faculdade de Filosofia, Ciências e Letras, Boletim 128, Filologia românica No. 3.) Pp. 71. São Paulo, 1951.
- Englisches Handwörterbuch ... By M. M. Arnold Schröer, ed. by P. L. Jaeger. Fasc. 9, *ha—indue*, pp. 625-720. Heidelberg: Carl Winter, Universitätsverlag, n.d.

- Die Entwicklung neuer germanischer Kultursprachen von 1800 bis 1950. By Heinz Kloss. (Schriftenreihe des Goethe-Instituts, Vol. 1.) Pp. 254. München: Pohl & Co., 1952.
- Esquisse d'une théorie des degrés de comparaison. By Roch Valin. (Cahiers de linguistique structural, publiés sous les auspices de la Faculté des Lettres de l'Université Laval, No. 2.) Pp. 20. Québec: Les Presses Universitaires Laval, 1952.
- Essai sur la phonétique du parler rhétoroman de la Vallée de Tavetsch (Canton des Grisons, Suisse). By Léonard Caduff. Pp. xiv, 229, offset. Bern: A. Francke AG, Verlag, 1952.
- Os estudos de filologia portuguesa de 1930 a 1949: Subsídios bibliográficos. By Giacinto Manuppella. (Instituto para a Alta Cultura: Publicações do Centro de Estudos Filológicos, No. 4.) Pp. 246. Lisboa: Centro de Estudos Filológicos, 1950.
- Études germaniques 7.81-240 (1952).
- Eusko-Jakintza 5.153-320 (1952).
- Faculty journal, No. 2 (1952). Education Faculty, Saga University, Japan.
- French influence in English phrasing. By A. A. Prins. Pp. vii, 320. Leiden: Universitaire Pers Leiden, 1952.
- Gace de la Buigne, Le roman des deduis: Édition critique d'après tous les manuscrits. By Åke Blomqvist. (Studia romanica holmiensia, Vol. 3.) Pp. 682, with 4 plates. Karlshamn [Sweden]: Almqvist & Wiksell (Stockholm), Librairie J. Thiebaud (Paris), 1951.
- Handbuch der Semasiologie: Kurze Einführung in die Geschichte, Problematik und Terminologie der Bedeutungslehre. By Heinz Kronasser. (Bibliothek der allgemeinen Sprachwissenschaft; erste Reihe: Handbücher.) Pp. 204. Heidelberg: Carl Winter, Universitätsverlag, 1952.
- Harvard journal of Asiatic studies 15.1-283 (1952).
- Hespéris 38.1-263 (1951).
- Hethitisches Wörterbuch: Kurzgefasste kritische Sammlung der Deutungen hethitischer Wörter. By Johannes Friedrich. (Indogermanische Bibliothek, 2. Reihe.) Fascicle 1, *a-to kalutiija-*, pp. 1-96. Heidelberg: Carl Winter, Universitätsverlag, 1952.
- L'histoire du morse. By V. Kiparsky. (Annales Academiae Scientiarum Fennicae, Series B, Vol. 73, No. 3.) Pp. 54. Helsinki: Suomalainen Tiedeakatemia, 1952.
- Humanisme en taalkunde. By L. L. Hammerich. (Scripta Academica Groningana: Aula-voordrachten, No. 2.) Pp. [vii], 46. Groningen, Djakarta: J. B. Wolters, 1952.
- 'Il est venu comme ambassadeur', 'Il agit en soldat' et locutions analogues en latin, français, italien et espagnol: Essai de syntaxe historique et comparée. By Veikko Väänänen. (Annales Academiae Scientiarum Fennicae, Ser. B, Vol. 73, No. 1.) Pp. 75. Helsinki, 1951.
- Die innere Form des Deutschen. By Hans Glinz. (Bibliotheca germanica, Vol. 4.) Pp. 504. Bern: A. Francke AG, Verlag, 1952.
- Institut Royal Colonial Belge 22.873-1132 (1951); 23.1-276 (1952).
- Internationale Zeitschriftenschau für Bibelwissenschaft und Grenzgebiete 1.i-xv, 1-196 (1951/52). Stuttgart: Verlag Katholisches Bibelwerk.
- International journal of American linguistics 18.115-278 (1952).
- Introductory Catalan grammar: With a brief outline of the language and literature, a selection from Catalan writers, and a vocabulary; 2nd edition, completely revised and enlarged. By Joan Gili. Pp. 189. New York: Hafner Publishing Co., 1952.
- Det isländska accenttecknet: En historisk-ortografisk studie (with an English summary). By Gustaf Lindblad. (Lunds Universitets Årsskrift, NF, Avd. 1, Bd. 48, Nr. 1.) Pp. 232. Lund: C. W. K. Gleerup, 1952.
- Italica 29.77-214 (1952).
- Izvestija Akademii Nauk SSSR, Otdelenie literatury iazyka 10.437-612 (1951); 11.1-278 (1952).
- Journal de la Société finno-ougrienne, Vols. 54 (1948-50), 55 (1951). Helsinki: Suomalais-Ugrilainen Seura.
- Journal of the Burma Research Society 34.1-72, i-vii (1951).
- Journal of the Linguistic Society of Japan, No. 21 (1952).
- Journal of the Oriental Institute 1.201-388, 17-32 (1952). M. S. University of Baroda, India.
- The journal of the Polynesian Society 60.93-269, Suppl. 1-145, 419-40 (1951).
- Die Landfahrzeuge des alten Mesopotamien. By Armas Salonen. (Annales Academiae Scientiarum Fennicae, Series B, Vol. 72, No. 3.) Pp. 197, with 50 plates. Helsinki, 1951.

- Language teaching in the Philippines: A report. By Clifford H. Prator Jr. Pp. 96. [Manila]: United States Educational Foundation in the Philippines, 1950.
- La langue est-elle ou n'est pas un système? By Gustave Guillaume. (Cahiers de linguistique structurale, publiés sous les auspices de la Faculté des Lettres de l'Université Laval, No. 1.) Pp. 30. Québec: Les Presses Universitaires Laval, 1952.
- Lateinisches etymologisches Wörterbuch. By Alois Walde, 3d ed. by J. B. Hofmann. (Indo-germanische Bibliothek; 2. Reihe: Wörterbücher.) Fascicle 18, *sēmita to stinguō*, pp. 513-92; 19, *stipa to terō*, pp. 593-672. Heidelberg: Carl Winter, Universitätsverlag, 1951, 1952.
- Leeds studies in English and kindred languages, Nos. 7-8 (1952). Leeds: Department of English Language and Medieval Literature, The University.
- Leuvense bijdragen, Vol. 41, Bijblad 73-141 (1951).
- Levende talen, Nos. 165, 166 (1952).
- Lingua nostra 13.29-96 (1952).
- Lulelappsk ordbok (Lulelappisches Wörterbuch). By Harald Grundström. (Skrifter utgivna genom Landsmåls- och Folkminnesarkivet i Uppsala, Ser. C.1.) Fascicle 10, *vuolatuhuttē to hōuri*, Tillägg, Person- och släktnamn, pp. 1441-587. Uppsala: A.-B. Lundequistska Bokhandeln; København: Einar Munksgaard, 1952.
- Lyydiläismurteiden äännehistoria, 2. Vokaalit. By Aimo Turunen. (Mémoires de la Société finno-ougrienne, No. 99.) Pp. viii, 266. Helsinki: Suomalais-Ugrilainen Seura, 1950.
- Le maître phonétique III.30.1-18, i-xi (1952).
- Malgache et maanjan: Une comparaison linguistique. By Otto Chr. Dahl. (Avhandlingar utgitt av Egede-Instituttet, Vol. 3.) Pp. 408. Oslo: Egede-Instituttet (Arne Gimnes Forlag), 1951.
- Man 52.65-144 (1952).
- Meaning, communication, and value. By Paul Keeskemati. Pp. viii, 349. Chicago: University of Chicago Press, 1952.
- Die mittelniederdeutsche Übersetzung der sog. Hieronymus-Briefe nach der Lübecker Handschrift (MS. Theol. Germ. 11). By Martta Jaatinen. (Annales Academiae Scientiarum Fennicae, Series B, Vol. 65, No. 2.) Pp. 136. Helsinki, 1950.
- Monotessaron: Eine mittelniederdeutsche, erweiterte Fassung vom Jahre 1513 (Diözesenarchiv, Trier, Nr. 75.) By Joh. Gerson, ed. by Axel Mante. (Lunder germanistische Forschungen, Nr. 25.) Pp. lxxii, 495. Lund: C. W. K. Gleerup; Kopenhagen: Ejnar Munksgaard, 1952.
- Morphologische Neurungen im altindischen Verbalsystem. By M. Leumann. (Mededelingen der Koninklijke nederlandse Akademie van Wetenschappen, Afd. Letterkunde, Nr. 15.3.) Pp. 51. Amsterdam: N. V. Noord-Hollandsche Uitgevers Maatschappij, 1952.
- Näytteitä äänis- ja keskivepään murteista. By E. N. Setälä and J. H. Kala, ed. by E. A. Tunkelo and Reino Peltola. (Mémoires de la Société finno-ougrienne, No. 100.) Pp. xvi, 621. Helsinki: Suomalais-Ugrilainen Seura, 1951.
- Nazva 'Ukraina' na Zakarpatti [The name 'Ukraine' in South Carpathia]. By B. Barvins'kyj. (Ukrainian Free Academy of Sciences; Onomastica, No. 4.) Pp. 16. Winnipeg: Nakladom Bratstva Karpats'kykh Sičovykiv, 1952.
- Neophilologus 36.65-192 (1952).
- Norsk tidsskrift for sprogvidenskap 16.1-495 (1952).
- Nueva revista de filología hispánica 6.1-108 (1952).
- Obras de Juan de Cueto y Mena: Edición crítica con introducción y notas. By Archer Woodford; preface by José Manuel Rivas Sacconi. (Publicaciones del Instituto Caro y Cuervo, Vol. 9.) Pp. xxxix, 315. Bogotá, 1952.
- Onseigaku-ronkoo [Treatise on phonology]. By Oonisi Masao. Pp. 337. Tokyo: Sinozaki-syorin, 1952.
- Orbis 1.1-327 (1952). Bulletin international de documentation linguistique. Louvain: Centre International de Dialectologie Générale.
- Orientalia NS 21.265-528 (1952).
- Our storehouse of Missouri place names. By Robert L. Ramsay. (University of Missouri bulletin, Vol. 53, No. 34; Arts and science series 1952, No. 7; Missouri handbook No. 2.) Pp. 160. [Columbia, Mo.]: University of Missouri, 1952.

- Le pélasgique: Essai sur une langue indo-européenne préhellénique. By A. J. van Windekens. (Université de Louvain, Institut Orientaliste: Bibliothèque de *Muséon*, Vol. 29.) Pp. xii, 179. Louvain: Publications Universitaires, Institut Orientaliste, 1952.
- Précis de sémantique française. By S. Ullman. (Bibliotheca romanica; series prima: Manualia et commentationes, No. 9.) Pp. [v], 334. Berne: Éditions A. Francke S. A., 1952.
- Recherches sur les mots minoens. By Const. D. Ktistopoulos. (Étude soumise à l'Académie d'Athènes.) Pp. 25, mimeographed. Psychiko (Athènes), 1952.
- Recuento de vocabulario español, Vol. 1. By Juana A. Méndez, directed by Ismael Rodríguez Bou. (Trabajos de investigación auspiciados por el Consejo Superior de Enseñanza: Publicaciones pedagógicas, Ser. 2, No. 12, Vol. 1.) Pp. xix, 668. Río Piedras: Universidad de Puerto Rico, 1952.
- Reflexive und intransitive Verba im älteren Westgermanischen. By Lars Hermodsson. Pp. 347. Uppsala: Almqvist & Wiksells Boktryckeri AB, 1952.
- Remarques sur la place du verbe dans la phrase active et moyenne en langue sanscrite. By J. Gonda. Pp. 86. Utrecht: A. Oosthoek, 1952.
- Revista brasileira de geografia 13.1-168 (1951).
- Revista de filología española 36.1-208 (1952).
- Revista del Museo e Instituto arqueológico, Nos. 13, 14 (1951). Universidad Nacional de Cuzco (Peru).
- Sanskrit sandhi and exercises. By M. B. Emeneau. Pp. [iii], 28. Berkeley and Los Angeles: University of California Press, 1952.
- Slavica canadiana A.D. 1951: A selected bibliography of Slavic books and pamphlets published in or relating to Canada. By Jaroslav B. Rudnyč'kyj. (Slavistica, No. 15.) Pp. 16. Winnipeg: Ukrainian Free Academy of Sciences, 1952.
- Smithsonian Institution, Bureau of American Ethnology, Bull. 147: Journal of an expedition to the Mauvais Terres and the Upper Missouri in 1850, by Thaddeus A. Culbertson, ed. by John Francis McDermott; pp. viii, 164, with 2 maps.—Bull. 148: Arapaho child life and its cultural background, by Sister M. Inez Hilger; pp. xv, 253, with 40 plates.—Bull. 150: The modal personality structure of the Tuscarora Indians as revealed by the Rorschach test, by Anthony F. C. Wallace; pp. viii, 120, with 1 plate. Washington: U. S. Government Printing Office, 1952.
- Smithsonian Institution, Institute of Social Anthropology, Publication No. 13: The Tajin Totonac, Part 1: History, subsistence, shelter and technology, by Isabel Kelly and Angel Palerm; pp. xiv, 369, with 33 plates. — Publ. 14: The Indian caste of Peru, 1795-1940: A population study based upon tax records and census reports, by George Kubler; pp. vi, 71. Washington: U. S. Government Printing Office, 1952.
- Smithsonian miscellaneous collections, Vol. 117, No. 12: Two aboriginal works of art from the Veracruz coast, by Philip Drucker (Publ. 4091); pp. 7, with 3 plates. Washington: Smithsonian Institution, 1952.
- Speculum 27.281-600 (1952).
- Språkvetenskapliga Sällskapets i Uppsala Förhandlingar Jan. 1949 - Dec. 1951. (Uppsala Universitets Årsskrift 1951, No. 9.) Pp. vii, 174. Uppsala: A.-B. Lundequistska Bokhandeln; Wiesbaden: Otto Harrassowitz, 1951.
- Studia linguistica 5.53-112 (1951).
- Studies in honor of Albert Morey Sturtevant. [Ed. by L. R. Lind.] Pp. [v], 169. Lawrence: University of Kansas Press, 1952.
- Taal en versbouw: Rede uitgesproken ... aan de Rijksuniversiteit te Utrecht op 19 Mei 1952. By A. Teeuw. Pp. 27. Djambatan-Amsterdam, 1952.
- Tau oder tau-t-an und das Rätsel der sprachlichen und menschlichen Einheit. By Gustav Zollinger. Pp. 98. Bern: A. Francke AG, Verlag, 1952.
- Les temps du verbe fini (indicatif) en français moderne. By H. Sten. (Det Kongelige Danske Videnskabernes Selskab, Historisk-filologiske Meddelelser, Vol. 33, Nr. 3.) Pp. 264. København: Ejnar Munksgaard, 1952.
- Thesaurus: Boletín del Instituto Caro y Cuervo 7.1-503 (1951).
- Tijdschrift voor nederlandse taal- en letterkunde 70.1-160 (1952).

- A unidade da România Ocidental. By Theodoro Henrique Maurer Jr. (Universidade de São Paulo: Faculdade de Filosofia, Ciências e Letras: Boletim 126, Filologia românica No. 2.) Pp. 227. São Paulo, 1951.
- Unghoswe waanyanja. By Bennett E. Malekebu, ed. by Guy Atkins. (Annotated African texts, No. 1: Maŋanja.) Pp. 124. Cape Town and London: Oxford University Press, 1952.
- Virittäjä 56.77-240 (1952).
- Virittäjän sisällys: Vuosina 1883, 1886 ja 1897-1946. By Erkki Itkonen and Pertti Virtaranta. Pp. 395. Helsinki: Suomalaisen Kirjallisuuden Seuran, 1952.
- Vokalismen i Iddemålet. By Reidar Myhre. (Skrifter frå Norsk Målførearkiv, No. 1.) Pp. iv, 118. Oslo: Jacob Dybwad, 1952.
- Volksbräuche und Volksdichtung der Wotjaken: Aus dem Nachlasse von Bernhard Munkácsi. Ed. by D. R. Fuchs. (Mémoires de la Société finno-ougrienne, No. 102.) Pp. xxxvii, 715. Helsinki: Suomalais-Ugrilainen Seura, 1952.
- William Barnes, linguist. By Willis D. Jacobs. (University of New Mexico publications in language and literature, No. 9.) Pp. 87. Albuquerque: University of New Mexico Press, 1952.
- Wogulische Volksdichtung. Gesammelt und übersetzt von Artturi Kannisto, bearbeitet und herausgegeben von Matti Liimola. (Mémoires de la Société finno-ougrienne, No. 101.) 1. Band: Texte mythologischen Inhalts, pp. xlii, 483. Helsinki: Suomalais-Ugrilainen Seura, 1951.
- Words and ways of American English. By Thomas Pyles. Pp. ix, 310. New York: Random House, 1952.
- Základná jazykovedná terminológia. [By Štefan Peciar and others.] (Odborná terminológia, Sväzok 1.) Pp. 59. Bratislava: Slovenskej Akadémie Vied a Umení, 1952.
- Zeitschrift für Phonetik und allgemeine Sprachwissenschaft 6.1-268 (1952).
- Zur syntax des tschechischen konjunktivs mit einem anhang über den russischen konjunktiv. By Gunnar Bech. (Travaux du Cercle linguistique de Copenhague, Vol. 7.) Pp. 131. Copenhagen: Nordisk Sprog- og Kulturforlag, 1951.

30 November 1952

THE BASIS OF GLOTTOCHRONOLOGY

ROBERT B. LEES

University of Chicago

[It is shown that a linguistic dating system can be set up on the basis of several explicit assumptions about morpheme decay. Thirteen sets of data, presented in partial justification of these assumptions, serve as a basis for calculating a universal constant to express the average rate of retention \bar{k} of the basic-root-morphemes: $k = 0.8048 \pm 0.0176$ per millennium, with a confidence limit of 90%. Finally an expression is derived for the sampling-error to be expected in the calculated time-depths of related dialects.]

0. INTRODUCTORY

0.1. Glottochronology. Many linguists have shown an increasing interest recently in the application of mathematical and, in particular, statistical methods to linguistic problems.¹ One such application is the use of statistical techniques to measure the degree of relatedness among cognate dialects.² It was in this connection that Swadesh first suggested in his article on Salish internal relationships the particular statistical method which has since been elaborated under the name GLOTTOCHRONOLOGY. The validity of the glottochronologic technique rests at present on the data and the mathematical derivation to be set forth in this paper.

0.2. Method. If (1) the morpheme inventory of a language, or a definable portion of it, is observed over a span of time, and if (2) the individual members of the inventory at a given time are identified as cognates of members at some previous time, and if (3) some statable regularity can be found in the time-rate at which members disappear from the inventory to be replaced by new items, then the number of items in a certain subset which are present at any one time can be used as a measure of time elapsed since some previous time-point for which a similar count is available. The members of the chosen subset may be likened

¹ See B. Trnka, *A tentative bibliography* (CIPL: Publication of the Committee on Linguistic Statistics, No. 1; 1950); H. E. Driver, *Statistics in anthropology*, *Am. anthr.* 55.4 ff. (1953), with bibliography.

² A. L. Kroeber and C. D. Chrétien, *Quantitative classification of Indo-European languages*, *Lg.* 13.83-103 (1937); id., *The statistical technique and Hittite*, *Lg.* 15.69-71 (1939); C. D. Chrétien, *The quantitative method for determining linguistic relationships*, *Univ. Cal. Pub. Ling.* 1.11 ff. (1948); M. Swadesh, *Salish Internal relationships*, *IJAL* 16.157 ff. (1950); id., *Lexico-statistic dating of prehistoric ethnic contacts*, *Proc. Am. Phil. Soc.* 96.457 ff. (1952); id., *Mosan I: A problem in remote common origin*, *IJAL* 19.26 ff. (1953); D. W. Reed and J. L. Spicer, *Correlation methods of comparing idiolects in a transition area*, *Lg.* 28.348 ff. (1952); R. B. Lees, *Relationships among the Germanic dialects by a statistical method* (unpublished, 1950); id., *The genetic relation in linguistics* (unpublished, 1952); C. F. Hockett, *Linguistic time perspective and its anthropological uses* (paper read before the summer meeting of the Linguistic Society, Bloomington, 1952); D. Taylor, *Sameness and difference in two Island Carib dialects*, *IJAL* 18.223 ff. (1952); E. Cross, *Chronometric and telemetric determination of the relationship of Latin, Gothic, Old Church Slavic, and their present affiliates* (paper read before the meeting of the Linguistic Society, Cambridge, 1952). This list is not exhaustive.

to the (indistinguishable) atoms in a given mass of a radioactive element; since the rate of disintegration is predictable at any time during observation of the sample, the mass (or number of remaining atoms) of this element remaining among the decay products at any time in the sample is a measure of how long the sample has been decaying. The analysis of decay products in mineral samples permits the calculation of the age of the earth's crust. Similarly, analyses of morpheme decay products should provide an absolute chronology for lexical history.

0.21. Morpheme decay. The total morpheme inventory of a language may contain over 100,000 identifiable items; of these, some smaller number, say 20,000 morphemes, comprises (according to most accepted estimates) the active colloquial vocabulary of an average speaker. Of the active colloquial morphemes, a few hundred may be affixes or other patterned units in a tight grammatical structure; these we shall disregard. The rest are active colloquial root-morphemes, each correlated with a small set of extralinguistic items in the culture, i.e. each with its own set of meanings.

These meanings may be of the most varied sort, differing from culture to culture; but certain meanings are likely to be found in all cultures, such as body parts, numerals, geographical terms, and simple activities.³ We call the morphemes associated with these cultural universals the BASIC-ROOT-MORPHEME INVENTORY, and denote this set of several thousand items by *I*. Because of the universality of the meanings associated with the members of *I*, we may assume that the size of *I* remains nearly constant in time and from language to language.

If the morphemes correlated with a certain subset of cultural universals in some language at a given time is compared with the corresponding morphemes correlated with the same meanings in the derivative cognate language at some later time, many corresponding morphemes will be found to be cognate; but a certain number may not be cognate.⁴ In the latter case, certain morphemes of the original set have disappeared and have been replaced by new, non-cognate morphemes. This temporal decrease in the size of the original subset is called MORPHEME DECAY.

The reasons for morpheme decay, i.e. for change in vocabulary, have been classified by many authors;⁵ they include such processes as word tabu, phonemic confusion of etymologically distinct items close in meaning, change in material culture with loss of obsolete terms, rise of witty terms or slang, adoption of prestige forms from a superstratum language, and various gradual semantic shifts, such as specialization, generalization, and pejoration.

³ Probably no meaning is completely universal: it would not surprise us to find languages containing no morpheme correlated with *man*, *head*, *three*, or *sky*, or languages in which *water* is represented by an affix (already excluded from consideration). We hope that such cases are so rare and so randomly distributed that our sampling process will show no significant effects from the non-universality of any one meaning.

⁴ By cognate we mean, of course, derivable one from the other by the use of a systematic set of phoneme correspondences, furnished by the traditional comparative method as applied to the language family in question.

⁵ See L. Bloomfield, *Language* 392-495 (New York, 1933); E. H. Sturtevant, *Introduction to linguistic science* 123-53 (New Haven, 1947).

0.22. Word-lists. In order to keep the material down to manageable size, and to render the material in one language comparable to that in another, the basic-root-morpheme inventory is sampled in the following way. A small set of basic morphemes (say 200) is selected from the inventory of some control language (say English), and each item in it is translated into the common colloquial expression of the test languages. These translations will then comprise, for the most part, root-morphemes which can be compared by the usual etymological techniques. Corresponding terms in two test languages will be either cognate or non-cognate, the latter label including terms borrowed by one language from the other. It is assumed that all the various causes of morpheme decay add up in both languages to some total amount of change which is dependent only upon the length of time during which these causes have been active.

Sometimes it will be difficult to determine whether two corresponding forms should be considered cognate root-morphemes. For example, the later form may be a fusion of several etymologically different roots; or it may represent either the retention of an old root or, just as plausibly, the borrowing of a similar root from some neighboring cognate dialect. The glottochronologic method depends upon the determination of retained original roots, and all ambiguous cases such as these are therefore ignored. It has been found that the number of items that must be omitted for this reason from a list of 215 basic-root-morphemes seldom runs higher than ten or fifteen.

1. RATE EQUATIONS

1.1. Morpheme decay law. We have described the basic-root-morpheme inventory I of a language as some large set of morphemes expected to remain approximately constant in size for all times and all languages. The identity of its members is continually changing as morphemes are lost and replaced by new, non-cognate morphemes. Let V be the number of items in I , and let R be the time-rate at which these items are exchanged.

As a BASIC WORKING HYPOTHESIS we shall assume that THIS RATE R IS ALSO CONSTANT IN TIME. This assumption will have to be verified by the data to be collected for calculating the value of the rate-constant (§3.2).

Consider some arbitrarily chosen sample subset S of the V items in I for a given language, containing N_0 items at the time t_0 , of which only N items are left at some later time t . Now while in one unit of time, R items have been replaced in I , only $(N/V)R$ items are lost from S ; for we may expect the random sample to lose the same FRACTION of items as the total population does in one unit of time. This is the rate at which S is decreasing in size, or:

$$-\frac{dN}{dt} = \frac{N}{V} R = \lambda N, \quad (1)$$

where λ is the rate-constant, and $\lambda = R/V$. Solving for N as a function of time:

$$N = N_0 e^{-\lambda t}, \quad (2)$$

where e is the well-known mathematical constant = 2.718. This is the BASIC RATE EQUATION for morpheme decay, and is identical in form to that for simple

radioactive decay. If we can evaluate the rate-constant λ , Eq. 2 tells us how many morphemes in a sample S are left, out of some original number N_0 , after time t has elapsed.⁶

Conversely, solving Eq. 2 for t gives us:

$$t = -\frac{1}{\lambda} \ln \frac{N}{N_0}. \quad (3)$$

This permits us to calculate how long it has taken a language with a morpheme decay rate-constant λ to decrease the size of a sample S of basic-root-morphemes from N_0 to N . We may call this value of t the **TIME-DEPTH** of the dialect stage which still had all the N_0 items, if we count the number N which are left from the present time (say 1953).

1.2. Time-depth equation. Suppose now that there are two genetically related dialects at time t , both assumed to have descended from some single common ancestor. Glottochronologically interpreted, this means that if we examine the identity of the items in a sample S of basic-root-morphemes in the two dialects back through time, there is a time t_0 when the two become identically the original N_0 items of the proto-language.

Throughout the time interval $t - t_0$ the two dialects have been diverging. By divergence we mean here that the particular items selected for replacement in S by new, non-cognate forms have not been the same for the two dialects. Some items will still be shared by the two, some will have been lost by one dialect, some by the other, and some by both.

1.21. Independence. By introducing two more fundamental hypotheses we shall be able to use these data (on the number of shared items left) to measure the time during which the two dialects have been diverging. First, as a **SIMPLIFYING ASSUMPTION** (which unfortunately it will prove difficult to verify), we shall suppose that **THE SELECTION OF ITEMS FOR REPLACEMENT IN ONE DIALECT IS STATISTICALLY INDEPENDENT OF THE SELECTION IN THE OTHER DIALECT**; that is, that the two selections are random with respect to one another. This may be stated mathematically in the following way: dialect A preserves N_a items out of the original N_0 after time t ; it therefore preserves the same fraction N_a/N_0 out of any random subset of S —in particular, the subset of N_b items retained by dialect B ; i.e. the number of items retained by A among those also retained by B is $N_s = (N_a/N_0)N_b$. Finally, the **FRACTION** of the N_0 original items still **SHARED** by dialects A and B after time t is:

$$F_0 = \frac{N_a}{N_0} \cdot N_b \cdot \frac{1}{N_0} = \frac{N_a N_b}{N_0^2}. \quad (4)$$

1.22. Interlingually constant rate. The second hypothesis which we shall introduce here is one that we shall later attempt to verify empirically: namely, that the rate-constant for morpheme decay has the same value for both dialects A and B : i.e. $\lambda_a = \lambda_b$.

1.23. Dating equation. Now since A and B have been diverging for the same

⁶ We shall even try to show later that the rate-constant may be considered the same for all languages.

length of time, and since we have assumed that $\lambda_a = \lambda_b$, we see from Eq. 2 that $N_a = N_b = N$, and we may rewrite Eq. 4 to read:

$$F_s = \left(\frac{N}{N_0} \right)^2. \quad (5)$$

Substituting in Eq. 2, we obtain:

$$F_s = e^{-2\lambda t}. \quad (6)$$

Eq. 6 enables us to calculate the fraction of items in S shared by the two dialects after time t , provided we know the value of the rate-constant λ .

As before, we may solve Eq. 6 for t to obtain the time-depth equation for two cognate dialects:

$$t = -\frac{1}{2\lambda} \ln F_s. \quad (7)$$

It will prove convenient to express the rate-constant in Eq. 7 somewhat differently. We shall determine empirically the average fraction of retained items per 1000 years. Measuring time in millenia, we may write this from Eq. 2:

$$\frac{N_1}{N_0} = e^{-\lambda} = k. \quad (8)$$

Then substituting our new constant k for λ in Eq. 7:

$$t = \frac{\ln F_s}{2 \ln k} = \frac{\log F_s}{2 \log k}. \quad (9)$$

2. EMPIRICAL DATA

In order to evaluate the rate-constant λ , or its equivalent, k , counts were made on a sample S of basic-root-morphemes for a number of languages to determine the fraction of the original N_0 items retained after the language had evolved through a time t . These data were then expressed as fraction-retained-per-millennium, or k , and an arithmetic mean was calculated. The data were also subjected to several statistical treatments to attempt a verification of our hypotheses.

2.1. Word lists. The basic-root-morpheme sample for each language tested was obtained by translating each of 215 English words into the most common colloquial term of that language. The first studies of this kind were made with the list used by Swadesh⁸ to measure the rate-constant of English. Since there

⁷ We use \ln for logarithms to the base e , \log for base 10. Eq. 9 is identical with Swadesh's equation

$$i = \frac{\log C}{2 \log r}$$

as given in *IJAL* 16.161 (1950).

⁸ Ibid. 161 (§2.2, Par. 1 and 3). The English words of Swadesh's list were originally chosen to be representative of universal semantic areas, to be relatively stable and resistant to culture changes, and to be easily found in the lexicons of many languages. Swadesh has since suggested a slightly modified word list, *Proc. Am. Phil. Soc.* 96.456-7 (1952).

is no a-priori reason to prefer one set of basic English morphemes to another, and since altering the word list might possibly introduce an unknown variable, all the data published here were obtained with the same list. The effects of changing the word list will have to be studied in the future.⁹

For each language tested, one list was prepared for an older stage and another for a more recent stage; in each case there was an independent way of dating the vocabulary of both stages. Corresponding morphemes were then compared by specialists in the language family involved, and word pairs were marked as cognate, non-cognate, or indeterminate. Occasionally it was impossible to translate an item, especially for the older stage; such items were simply omitted. Uncertainties and omissions never reduced the total of any list to fewer than 200 items.

2.2. Data. At present, 13 word counts have been prepared for this study; the following list shows the languages compared, the number of words and cognates (with percent of cognates in parentheses), and the persons responsible for the test.

1. Old English of 900-1000 A.D. : Modern English; 209 words, 160 cognates (76.6); R. B. Lees and J. H. Sledd. — 2. Plautine Latin of 200 B.C. : early Modern Spanish of 1600 A.D.; 200 words, 131 cognates (65.5); D. Griffin. — 3. Plautine Latin : Molière's French of 1650 A.D.; 200 words, 125 cognates (62.5); D. Griffin. — 4. Old High German of 800-900 A.D. : Modern German; 214 words, 180 cognates (84.2); R. B. Lees and G. J. Metcalf. — 5. Middle Egyptian of 2100-1700 B.C. : Coptic of 300 A.D.; 200 words, 106 cognates (53.0); K. Baer. — 6. Koine Greek of 250 B.C. : Modern Athenian Greek; 213 words, 147 cognates (69.0); E. P. Hamp and B. Einarson. — 7. Koine Greek : Modern Cypriote; 211 words, 143 cognates (67.8); E. P. Hamp and B. Einarson. — 8. Ancient Classical Chinese of 950 A.D. : Modern Mandarin; 210 words, 167 cognates (79.6); C. Y. Fang, M. Swadesh. — 9. Old Norse of 800-1050 A.D. : Modern Swedish; 207 words, 176 cognates (85.0); R. B. Lees and G. Franzen. — 10. Classical Latin of 200 B.C. : Modern Tuscan; 210 words, 144 cognates (68.6); J. Corominas and H. Noce. — 11. Classical Latin : Modern Portuguese; 210 words, 132 cognates (62.9); J. Corominas. — 12. Classical Latin : Modern Rumanian; 209 words, 117 cognates (56.0); J. Corominas. — 13. Classical Latin : Modern Catalan; 208 words, 126 cognates (60.6); J. Corominas.

2.3. Rate-constants. These data can now be tabulated, together with the calculated rate-constants and time-depths (counted from the mid-points of the ranges in dates). The rate-constant k is fraction-retained-per-millennium, and t is given in millenia.

LANGUAGE	F_s	t	k
1. English	.766	1.0	.766
2. Spanish	.655	1.8	.790
3. French	.625	1.85	.776
4. German	.842	1.1	.854

⁹ To save space we will not list the 215 words used, but a copy of the work-sheet may be obtained from the author upon request. All the original data are available for inspection.

LANGUAGE	F	t	k
5. Coptic	.530	2.20	.760
6. Athenian	.690	2.07	.836
7. Cypriote	.678	2.07	.829
8. Chinese	.796	1.0	.795
9. Swedish	.850	1.02	.854
10. Italian	.686	2.15	.839
11. Portuguese	.629	2.15	.806
12. Rumanian	.560	2.15	.764
13. Catalan	.606	2.15	.793

$$\text{Mean } \bar{k} = .8048 \pm .0176/\text{mill.}$$

The limits of error on the mean rate-constant \bar{k} were calculated as the 9/10-error in the mean, using small-sample methods.¹⁰

We take this to mean that on the average about 81 % of the basic-root-morphemes of a language will survive as cognates after 1000 years, for all languages, at all times.

3. VERIFICATION OF HYPOTHESES

We must now see if these empirical data can be used not only to provide a value for an average rate-constant, but also to verify the hypotheses proposed for justifying the calculation of a mean rate-constant for all languages.

3.1. Interlingually constant k . The theory mentioned in §1.22 assumes that the rate-constant λ for morpheme decay is the same in two dialects. The obtained values of k ($= e^{-\lambda}$, Eq. 8) range from 0.760 for Coptic : Egyptian up to 0.854 for both German : Old High German and Swedish : Old Norse, with a standard-deviation of 0.0342.¹¹ To what extent are we justified in assuming that all of this variation is random sampling error, and that, if it were not for these supposedly random perturbations, each of the thirteen languages would show the same rate-constant?

3.11. The χ^2 -function. To answer this question it is not sufficient to look at the differences between the expected and the observed result ($E - O$), for there is

¹⁰ That is,

$$0.0176 = \frac{z\sigma}{\sqrt{n-1}},$$

where σ is the standard-deviation of the k 's by the small-sample formula

$$\sigma = \frac{\sum \delta^2}{\sqrt{n-1}},$$

$n = 13$ is the number of k 's averaged, and z is the unit-deviate for 12 degrees of freedom on the 't'-distribution for the 90% level of confidence. We have used the 9/10-error as a measure of uncertainty throughout this study.

¹¹ The standard-deviation is the root-mean-square deviation about the mean:

$$\sigma = \sqrt{\frac{\sum_i (k_i - \bar{k})^2}{n-1}}.$$

no way to evaluate the significance of a larger or smaller difference. To be sure, there will always be a discrepancy; we seek merely to know what maximum discrepancy may be tolerated in any calculations that use our data. To evaluate the discrepancy between expected and observed results we may employ a certain function of the differences ($E - O$), a function whose mathematical and statistical properties have been studied and tabulated.¹²

This function is known as CHI-SQUARE, and it is given by:

$$\chi^2 = \sum_i \frac{(E_i - O_i)^2}{E_i} \quad (10)$$

where the O_i are a set of observations on the outcome of some experiment, and the E_i are the expected values as predicted by some theory. For every number of independent observations, and every value of χ^2 for them, there is a certain known probability that a value of χ^2 as great or greater could have been obtained by chance alone. If our calculated value of χ^2 is small enough that this probability P is high (say 50%), then we have reason to believe that the discrepancy between E_i and O_i was due only to random factors, to be expected in any physical data. But if our value of χ^2 is so large that P is very low (say 1% or less), then the discrepancy probably did not result from chance alone in our sample experiments, but represents a 'real' difference between observation and theory.

Now our theory predicts the outcome of thirteen probability experiments, wherein N_0 morphemes are allowed in each case to fall by chance into two categories, RETAINED and LOST. It tells us that for a given time t , the number retained is

$$N = N_0 e^{-\lambda t}$$

and the number lost is

$$N_0 - N = N_0(1 - e^{-\lambda t}).$$

The actually observed outcomes for the 'retained' category are found in the table in §2.2, and the number in the 'lost' category is obtained in each case by subtraction. There are then 26 comparisons of observation and theory from which a calculation of χ^2 may be made. The value of χ^2 was computed for our thirteen sets of data:

$$\chi^2 = 29.5.$$

Finally, from statistical tables of the χ^2 function we see that for 12 degrees of freedom¹³ there is a probability of 0.01 for obtaining a χ^2 at least as great as

¹² The author gratefully acknowledges the invaluable help that he received, in preparing this formulation, from W. Kruskal and his associates in the Department of Statistics, University of Chicago.

¹³ The number of INDEPENDENT observations used to calculate χ^2 , however, is only 12, because the 13 in the 'lost' category may be determined from N_0 and the value in the 'retained' category in each case, and one last value may be determined from the remaining 13 and their mean (or its equivalent, the mean k), leaving altogether only 12 DEGREES OF FREEDOM, not 26.

26.1. Our value is even larger. We must conclude then, that if the χ^2 -test is valid for these data, there is very probably a greater-than-chance discrepancy between our theory and our data.

Before accepting this conclusion, however, we must note several important qualifications. First of all, the value of χ^2 depends upon the size of E_i and O_i ; in our case these numbers are rather large, in the neighborhood of 100 (40 to 170). This fact may account to some extent for the high value of χ^2 .

Second, some of the 'real' discrepancy between E_i and O_i may be contained in the values of the parameter t ; surely we cannot claim to have dated all the word lists very accurately. Slight errors in dating our test lists will show up as errors in the predicted values E_i , which will not be due to a failure of our theory.

Third, we must remember that the χ^2 -test is by no means the only judge of a good fit; even with a high χ^2 , a theory may, within quite respectable limits, be a valuable tool for certain calculations.

Finally, there is some possibility that the value of \bar{k} , the mean rate-constant used to calculate the E_i , was faulty. We obtained it by averaging the values of k ; perhaps a better value (say, a maximum-likelihood value) would have been some other function of the individual k 's, say $e^{-\lambda}$. (The equations for a minimum- χ^2 and for a maximum-likelihood calculation were both too difficult to solve analytically.)

3.12. The 9/10-error. Another indication of the reliability of our assumption of interlingually constant k is the sampling-error in \bar{k} itself. We have calculated the standard-error of the k 's by small-sample methods:

$$\sigma_{\bar{k}} = \frac{1}{n-1} \sqrt{\sum \delta^2} = 0.00987$$

and the 9/10-error in the mean:

$$9/10\text{-error } (\bar{k}) = dk = z_{13}\sigma_{\bar{k}} = \pm 1.782 \sigma_{\bar{k}} = \pm 0.0176$$

where z_{13} is the unit deviate at $P = 0.1$ for 13 observations on the ' t '-distribution. This number, the 9/10-error, means the following: if the fluctuations among the values of k are truly random, and if we draw a large number of such samples of 13 from among the whole population of possible languages, and compute similar sample means \bar{k} for each, then on the average 90 % of these sample means will lie within 0.0176 on either side of the mean we obtained, viz. $\bar{k} = 0.8048$.

Now if 90 % of all sample means can be expected to lie within 2.2 % of our mean \bar{k} , we seem justified in assuming some considerable central tendency among the k 's. It is quite certain that they are NOT all identical, and indeed, statistical theory would predict just this fact; but all we desire is that the k 's be sufficiently close together so that our assumption that they are all equal will not introduce an intolerable error into further calculations.

3.2. Temporally constant k . In §1.1 we stated our first working hypothesis, that the rate-constant R (or λ , or k) is constant in time for any language. In order to check this assumption it would be necessary to measure the rate-constant for a given language at various periods in its history. Since we are reluctant to accept any data for time-depths of less than 500 years, it is very

difficult to find a language for which word lists could be prepared at 500-year intervals over several thousand years.

Perhaps the only language for which this could be done is Assyro-Babylonian, which covers about 5000 years of written records. But it is difficult to obtain word lists by the chosen method from existing Assyriological materials, and the author has not yet been able to assemble the necessary data.¹⁴

Some evidence tending to justify the assumption of a temporally constant k is seen among the thirteen determinations of §2.3, where we can find no discernable correlation between t and k .

3.3. Independence of selection. Our other basic assumption (in §1.21) is that the selection of items for replacement in the basic-root-morpheme inventory of one dialect is statistically independent of the selection in another dialect. Now we know of course that for two diverging dialects there can be no complete statistical independence, for the internal forces (whatever they are) which originally caused morpheme decay in the undifferentiated proto-language will certainly continue to affect the two in the same way, at least initially. Furthermore, although the two speech communities are geographically separated, so long as they remain in communication we shall expect certain external factors as well to affect them equally.

The only convenient way to test this hypothesis is to calculate the time-depth of related dialects which began diverging at a known date. If we find any discrepancy in time-depth, we shall ascribe it to a lack of independence.

3.31. German/English time-depth. When we compare the word lists for Modern German and Modern English, we find 124 cognates (58.5%) in a total of 212 words. By means of Eq. 9 we can calculate the time-depth:

$$t = \frac{\log .585}{2 \log .805} = 1.236 \text{ millennia.}$$

Counting 1,236 years back from 1952, we would predict that German and English began to diverge in basic-root-morpheme inventory about 716 A.D. But since the Germanic invasions of Britain began about 449 (though there was probably considerable traffic and intercommunication up to the year 600), our estimate would seem to be too late: the Middle German dialects which were the main source of Modern German must have separated from the northern dialects which were transplanted to Britain several centuries at least before our date.¹⁵

Before we ascribe this deviation to lack of independence between the two dialects, we must assess the limits of error in our answer to see if the allowable range does not perhaps include the historical date (see §4.3).

3.32. Turkish/Azerbaijani/Uzbek time-depths. A similar calculation was made

¹⁴ Members of the Assyriology staff in the Oriental Institute, University of Chicago, may be able to supply the necessary word lists.

¹⁵ One could object that part of this discrepancy may be due to failure of our first hypothesis of interlingually constant k . But if we recalculate, using the individual rate-constants for German and English instead of the mean constant k , we reduce the date by only 27 years, to 689 A.D.

for modern Istanbul Turkish and modern Azerbaijani.¹⁶ The word list contains 209 morphemes and 166 cognates (79.4%), giving a time-depth of 0.526 millenia. This would date the split of Osmanli and Azerbaijani about 1424 A.D. The Turks took Constantinople in 1453, but had been in Anatolia since about the year 1000. There must have been some considerable intercommunication between the Anatolian and the Caucasian Turks over a period of many years, and this may account for our late estimate.

A second word list was prepared for Turkish and for Ferghana Uzbek.¹⁷ On the basis of 177 usable words, of which 117 (66.2%) were identifiable cognates, the calculated time-depth was 0.954 millenia. This would indicate that the Osman tribes may have separated from their Uzbek speaking relatives about the year 1000, which compares favorably with their date of entry into Anatolia.

3.4. Miscellaneous calculations. Swadesh has calculated an Eskimo/Aleut time-depth and compared it with a carbon-14 dating of the earliest Aleutian settlement. Both dates were in the neighborhood of 3000 years.¹⁸

Taylor and Swadesh¹⁹ have counted morphemes retained in Modern Carib from the Dominican Carib of 1650 A.D.; they find 93.5% cognates left in the modern dialect. Using our mean rate-constant we would put the split at 1640.

The author has prepared a word-list for French and English to estimate the time-depth of Germanic/Romance. The list of 202 morphemes shows 56 cognates (27.7%), dating the separation of these two branches at 1000 B.C. Now Trager and Smith have proposed²⁰ that Southwest European may have split off from General European (to leave North European) between 1800 and 1500 B.C., in any case not much later. Our late date may very well reflect some contact between Romance (or Italic) and Germanic after that time, and between Norman French and Old or Middle English.

The author has also estimated the fraction of shared cognates for Modern Gujarati and Modern Rajasthani²¹ as 104/191 or 54.5%, which would put the split of these two Indic dialects at about 550 A.D.

A. Barrera Vasquez (Education Clearing House, Unesco), prepared word lists for the Mayan of Yucatan, using missionary materials from 1540-1700 and the modern language. There were 203 cognates out of 212 words (95.8%), giving a time-depth of 200 years. This is probably too short a time to insure any great precision for the estimate.

To observe the effect of including in the word list a large proportion of non-basic items, the entire English vocabulary of C. D. Buck's *Dictionary of selected*

¹⁶ The Azerbaijani word list was provided by F. W. Householder Jr., of Indiana University.

¹⁷ The Uzbek words, taken from the author's field notes, were supplied by Rusi Nasar of Margalan, Uzbekistan.

¹⁸ *Proc. Am. Phil. Soc.* 96.452-3 (1952). The difficulty with this example is that (1) we cannot be sure that the owners of the organic matter from which the C-14 was taken spoke Eskimo-Aleut, and (2) we do not know how long after or before the deposition of the sample the two branches of Eskaleut split apart.

¹⁹ Taylor, *IJAL* 18.229 ff. (1952).

²⁰ G. L. Trager and H. L. Smith Jr., *A chronology of Indo-Hittite*, *SIL* 8.61 ff. (1950).

²¹ Using as informants Nataraj Vashi of Bombay and Kumar Sumer Singh of Jaipur.

synonyms in the principal Indo-European languages was examined for cognates retained from Old English. The root-morphemes include many from such semantic areas as religion, government, and military life. Out of 1010 words used, 573 (56.8%) are cognate. Assuming a time-depth of 1000 years, this figure gives a retention rate of 0.568/millennium, nearly 30% lower than our mean rate-constant 0.805.

In Buck's chapter on body parts and functions alone, there are 68 cognates out of 97 words. The indicated rate of 0.701 for these items is still 13% below our mean. This may perhaps reflect a large incidence of tabu among the items of this semantic area.

3.5. Homogeneity. The randomness or representativeness of our word list was checked once by splitting the German/English list (§3.31) into two parts in a statistically arbitrary way, viz. alphabetically. The fraction of shared items for the two halves was calculated as follows: items 1-112, 65 cognates, $F_s = 0.580$; items 113-213, 59 cognates, $F_s = 0.590$. This close check would imply that the list used is at least homogeneous.

4. LIMITS OF ERROR

We come finally to the most important consideration of all. Assuming that the rate-constant for morpheme decay is real and has the value that we obtained for it (§2.3), and assuming further that it is interlingually and temporally constant, with what degree of precision does this method allow us to specify our data?

4.1. Errors in F_s and k . In §3.12 we showed that we can estimate the limits of sampling-error in our mean rate-constant by assuming that the differences among the individual k 's are random (and normally distributed). We expressed this as the 9/10-error:

$$dk = \pm 0.0176$$

Now the fraction of shared items F_s obtained from correlated word lists for two related dialects is a proportion drawn on a sample of some 200 items, presumably representative of the entire basic-root-morpheme inventory of each language. We shall accordingly expect to find also a sampling error in F_s . This may be expressed as a standard-error in the proportion F_s :

$$dF_s = \sqrt{\frac{F_s(1 - F_s)}{m}} \quad (11)$$

4.2. 9/10-error in t . Finally, since both of these sampling-errors will contribute to errors in t , we can calculate from them a 9/10-error in t :²²

$$dt = \pm 0.0946 \sqrt{t^2 + \frac{1569}{m} \cdot \frac{1 - F_s}{F_s}} \quad (12)$$

where m is the number of words in the list used.

²² From Eq. 9,

$$t = \frac{\ln F_s}{2 \ln k},$$

4.3. Range in time-depths. We may now return to the calculated time-depths in §3.3 and §3.4, and compute the limits of error in t .

The time-depth for English/German was given (§3.31) as 1.236 millenia. From Eq. 12 we calculate the 9/10-error in t :

$$dt = \pm 0.246 \text{ mill.},$$

and then we may state:

$$t = 1.236 \pm 0.246 \text{ mill. at the 90 \% level.}$$

On a calendar scale, German and English may be said to have separated lexically somewhere between 470 and 962 A.D., with 90 % assurance within the limitations of our theory. This means that if we recompute the time-depth for German/

and expanding in a Taylor's series about the point (k_0, F_0) :

$$t - t_0 = \frac{\partial t}{\partial k} (k - k_0) + \frac{\partial t}{\partial F} (F - F_0) + \dots,$$

where the partial derivatives are to be evaluated at the point (k_0, F_0) . Neglecting terms of higher order, multiplying out, and collecting constant terms:

$$t = \frac{\partial t}{\partial k} k + \frac{\partial t}{\partial F} F + \left(\frac{\partial t}{\partial k} k_0 + \frac{\partial t}{\partial F} F_0 + t_0 \right).$$

Now since $\text{var}(c) = 0$, $\text{var}(x + y) = \text{var}(x) + \text{var}(y)$, and $\text{var}(cx) = c^2 \text{var}(x)$, the variance in t is given by:

$$\text{var}(t) = \left(\frac{\partial t}{\partial k} \right)^2 \text{var}(k) + \left(\frac{\partial t}{\partial F} \right)^2 \text{var}(F),$$

or, since $\text{var}(x) = (\sigma_x)^2$,

$$\sigma_t = \sqrt{\left(\frac{\partial t}{\partial k} \right)^2 \sigma_k^2 + \left(\frac{\partial t}{\partial F} \right)^2 \sigma_F^2}.$$

Substituting for the partial derivatives

$$\frac{\partial t}{\partial k} = -\frac{t}{k \ln k} \quad \text{and} \quad \frac{\partial t}{\partial F} = \frac{1}{2F \ln k}$$

and also the expression for

$$\sigma_F^2 = \frac{F(1-F)}{m},$$

we obtain:

$$\sigma_t = \frac{\sigma_k}{k \ln k} \sqrt{t^2 + \frac{k}{4\sigma_k m} \cdot \frac{1-F}{F}}.$$

Substituting for the values $k = 0.8048$ and $\sigma_k = 0.00987$, and converting to the 9/10-error by $(9/10\text{-error}) = 1.645 \sigma$:

$$dt = \pm 0.0946 \sqrt{t^2 + \frac{1569}{m} \cdot \frac{1-F}{F}},$$

where m is the number of items in the word list.

English many times, with other word lists and mean rate-constants, 90 % of these dates will fall between 470 and 962.

Now this range may be reduced indefinitely by accepting lower and lower degrees of assurance. For example, we may say these two languages began diverging lexically between 666 and 766 A.D., but now with only 24 % assurance.

Similarly the other data in §3.32 and §3.4 may be expressed as follows, in each case with 90 % confidence:

Turkish/Azerbaijani	1277-1571 A.D.
Turkish/Uzbek	775-1217 A.D.
French/English	1509-491 B.C.
Gujerati/Rajasthani	258-836 A.D.
Maya of Yucatan	1678-1796 A.D.

4.4. Relative error. In §4.2 we showed how to estimate the 9/10-error in t . This number increases in absolute value from 0 years at $t = 0$, to ± 1000 years at $t = 5,600$ years, rising then ever more rapidly; at $t = 20,000$ years, the 9/10-error in t is equal in magnitude to t itself. We may also express the error as a fraction of t , i.e. as a RELATIVE-ERROR. When t is very small the relative-error $|dt|/t$ is large, dropping rapidly to a minimum of 17 % at $t = 3,200$ years and then rising again to 100 % at $t = 20,000$ years.

Taking our value of $\bar{k} = 0.8048$, and assuming $m = 215$ words, we may calculate the values of t , $|dt|$, and $|dt|/t$ corresponding to various values of F_s :

F_s	t	$ dt $	$ dt /t$	F_s	t	$ dt $	$ dt /t$
.99	23	26	115.0	.45	1,850	339	18.3
.95	119	61	50.9	.40	2,120	380	17.9
.90	244	90	36.8	.35	2,440	425	17.4
.85	377	115	30.5	.30	2,790	480	17.2
.80	517	139	26.9	.25	3,220	548	17.0
.75	666	163	24.5	.20	3,730	630	16.9
.70	826	193	23.4	.15	4,390	750	17.1
.65	1,000	215	21.4	.10	5,340	935	17.5
.60	1,182	241	20.4	.05	6,950	1,312	18.9
.55	1,385	270	19.5	.02	9,060	2,020	22.3
.50	1,608	302	18.8	.01	10,690	2,780	26.0

From this table of relative-errors we see that the optimum values appear between 1,000 and 10,000 years of time-depth, where $|dt|/t$ is less than 23 %. The relative-error in this time range could be reduced to less than 9.5 % if we would accept a 50 % confidence-level (sometimes called the probable-error), instead of 90 %.

4.5. Reduction of error. In this section we shall indicate briefly a few improvements which might be made to reduce the errors in the glottochronological method. By reduction of error we mean, of course, improvement of the precision (that is, the confidence) with which we may assert our predictions, not

improvement of the accuracy, which depends upon the quality of the data used and upon the theory itself.

4.51. Increase in the quantity of data. Since the sampling-error in a statistic varies inversely with the number of observations used in determining the value of the statistic, we can reduce the size of the error by using more data. By using twice as many words in our list (i.e. 430 instead of 215), we can reduce the error in t by at least 20-25 %, for now

$$dt = \pm .0946 \sqrt{t^2 + \frac{785}{m} \frac{1 - F_s}{F_s}}.$$

Similarly, by obtaining ten more determinations of k to include in our average \bar{k} , even though this causes no decrease in the dispersion of the k 's about their mean, we can still cut in half the sampling-error in \bar{k} .

The 9/10-error in t will then decrease by 30-35 %:

$$dt = \pm .0473 \sqrt{t^2 + \frac{3138}{m} \frac{1 - F_s}{F_s}}.$$

4.52. Improvement of the word list. It may be possible to define certain classes of morphemes which exhibit different and perhaps more stable rates of decay. For example, if our word lists consisted only of morphemes referring to body parts, the decay-rate-constants for the test languages might all group more tightly about their mean than do those which we have determined. They would then provide a better measure of time-depth, or at least a more precise measure, i.e. a measure with lower sampling-error.

If there are classes of basic-root-morphemes which decay at different rates, the assumption of a temporally constant k for our word list (which presumably contains members of several such classes) is clearly false. Through a given time period, as the more resistant classes survive the less resistant, their concentration in the sample will increase, and the value of k will rise.

OLD CHURCH SLAVONIC *bedrno*

HORACE G. LUNT II

Harvard University

[OCS **bedrno* 'rich' is attested only once, in the form *bedrno*. It has no relation to *bdr* 'vigilant' and *bđeti* 'be awake'. It is cognate with Sanskrit *bhadrā* 'good fortune, auspicious', and establishes an IE **bhed-r-* rather than the *bhād-* posited by earlier investigators. From the same base are formed Slavic names for the herb burnet saxifrage and the ladybug, both of which are associated in Slavic belief with good luck and prosperity.]

In discussing the so-called jer-umlaut in Old Church Slavonic, Diels cites *bedrno* (Euch. 58a3), which he connects, somewhat hesitantly, with *bdr* 'ready, willing, eager'. Vaillant makes the same connection without qualification.¹ The argument is that the root vowel *o* in the verb *bđeti* 'watch, be awake', became *o* under the influence of the front vowels (*ě, i*) which appear in the second syllable throughout the conjugation. The derivative adjective then became *bdr* by analogy, even when the jer was in strong position. OCS spellings seem to bear out this contention, for *o* is scarcely attested.² Russian *bodr*, however, assures the etymological *o*, and no one can object to the established connection of *bdr* and *bđeti* with Lith. *budrūs* 'watchful' and the Slavic verbs *bljud-* 'watch, observe' and *bud-* 'waken' and IE **bhudh-/bheudh-/bhoudh-*.³

A number of early texts (not OCS) have forms of *bdrno* 'watchful, vigilant';⁴ its derivation from *bdr* is clear from the meaning, the formation, and the usage in modern Russian Church Slavonic. There seems to be no modern evidence at all to support the presumed change *o* > *o* > *e* in this root.

In Euchologium, the prayer for a deceased priest asks that he receive *bedrno vzdanie* (sc. *vzdajanie*). It was the meaning 'watchful reward' which disturbed Diels, but matters were scarcely improved by Frček's discovery of the Greek original.⁵ The passage is slightly corrupt in the Slavic manuscript, but it is

¹ P. Diels, *Altkirchensl. Gram.* 110 (Heidelberg, 1932); A. Vaillant, *Manuel du vieux slave* 1.52 (Paris, 1948), and *Grammaire comparée des langues slaves* 1.135 (Paris, 1950).

² Only *Savvina Kniga* has *bdr*, *bđite* and *bđěti*, 3 *o* ~ 7 *o*; elsewhere only *o* is found in all possible forms. However, OR texts like the Ostromir and Archangel Gospels have *o* regularly.

³ See Berneker, *Slav. etym. Wörterbuch* s.v. *bđp*, *bljudp*, *budp* (Heidelberg, 1924). It should be noted that in Old Church Slavonic the adj. *bdr* is attested only in the identical phrases of Mt. 26.41 and Mk. 14.38, 'The spirit is WILLING/READY' (*πρόθυμος*). In later texts it is found meaning 'awake', but in modern Russian it has been extended to 'ready, forward, bold, cheerful'. From Russian it has been borrowed into literary Czech, Slovene, Bulgarian, and Macedonian.

⁴ Cf. e.g. Bologne Psalt., ps. 118.17, commentary; Sreznevskij, *Materialy dlja slovarja drevnerusskogo jazyka* s.v. *bdrno*, *bdrnyi*; the Yugoslav Academy's *Rječnik hrvatskoga ili srpskoga jezika* s.v. *bdrn*.

⁵ See J. Frček, *Euchologium Sinaiticum* 91 [776] [= *Patrologia orientalis* 24.5] (Paris, 1933). Since his OCS text is based on the old edition of Euch. by Geitler, generally agreed to be full of errors, Frček felt free to make emendations or at least suggestions on the basis of the Greek.

clear that the phrase in question translates *πλουσίαν τὴν ἀνταπόδοσιν* 'the RICH reward'. Frček rejects any connection with *bedrěnz* 'vigilant', and cites an Old Russian text where *bedri byša* translates *εὐθηνούσαν* 'prosper'.⁶ He sees here 'sûrement' a corruption with *bedri* for **gobъzъni*, a juxtaposition arrived at by comparison with the standard Slavonic *ugobzъša sja* in this passage. Nahtigal holds that the OR *bedri byša* is correct, but should be understood as 'tranquillus'; I must admit that I cannot follow his logic and have been unable to check the references he gives.⁷ His belief that the use of *bedri* to gloss *εὐθηνεῖν* forms a bridge to *bedrěno* 'rich' may be more acceptable.

Bedrěno means 'rich' in Euch. Let us suppose that this spelling represents *bedrěno*, which can be analyzed as a base *bedr-* plus the common suffix *-en-* plus the neuter *-o*. In Sanskrit there is a word *bhadrāḥ* meaning 'welfare, good fortune, prosperity, wealth; praiseworthy, good, auspicious'. The compound *sū-bhadrā-* is paralleled exactly by Avesta *hubaδra-* 'fortunate'. Pokorny follows Walde and Feist in deriving these forms from **bhād-* and connecting them with Gothic *batiza* 'better', *batista* 'best', OE *bet(e)ra*, *betst*, etc.⁸ There is nothing in the Germanic to forbid the assumption of a root **bhod-*, and the Indo-Iranian could come from either this form or the *e*-grade **bhed-*. Accepting the latter, one could posit an IE **bhedr-* 'prosperity, good fortune, wealth; prosperous, lucky', from which an OCS **bedrěnz* would be normal.

Perhaps there was also an adjective **bedrě* 'prosperous, fortunate', used by the South Slavic translator of Lamentations. In the process of copying this word, Russian scribes may have distorted it to the more familiar *bedri*; the meaning is changed but the passage still makes perfectly good sense. 'Her enemies are prosperous' becomes 'her enemies are bold' (or 'forward, cheerful').

The root *běd-* has been used to explain the name of the common herb burnet saxifrage (*pimpinella saxifraga*). Brückner cites only Polish *biedrzeniec* and *biedrznik*, and waves aside the problem of the vocalism. Vasmer notes Russian dialectal *bedreněc* and also Czech *bedrník*; he is somewhat reserved in noting that the 'vocalic assimilation' could have taken place only in Russian and that therefore the West Slavic words must be borrowings, with an 'incomprehensible' change in suffix.⁹ However, Russian also has *bedreník*, Ukrainian *bedryněc* and *bedrýč*, Byelorussian *bedrínec*; and we find *bedr'nac* and *będrnik* in Slovene, *bědrinac* and *bedrenik* in Serbocroatian¹⁰.

⁶ The source is Sreznevskij, *Materialy* s.v. *bedryi*, and the text adduced is Lamentations 1.5. It is not noted in this entry whether, as Nahtigal assumes, this citation refers to the famous 15th-century text including the copy of the colophon of the scribe Upyr' Lixyi and the date 1047; Sreznevskij gives no date for the other ms. of Lam. which he excerpted.

⁷ R. Nahtigal, *Euchologium Sinaiticum* 2.146 = *Dela* 2 of the Akademija znanosti in umetnosti (Ljubljana, 1942). For 'tranquillus' he refers to 'Concordantiae graecae in Septuaginta interpretes I 663, Amstelodami etc. 1718; Sveto pismo IV 342, v Ljubljani 1859', neither of which was accessible to me.

⁸ J. Pokorny, *Indogermanisches etym. Wörterbuch* s.v. *bhād*, gives the chief references. The modern Iranian languages, as R. N. Frye has kindly informed me, have retained the root, as in New Persian *bahr* 'beauty; fortunate'. He also points out the Aramaic *bahray*, borrowed from Middle Persian.

⁹ A. Brückner, *Słownik etymologiczny języka polskiego* s.v. *biedrzeniec* (Cracow, 1927); Max Vasmer, *Russisches etymologisches Wörterbuch* s.v. *bedreněc* (Heidelberg, 1950).

¹⁰ V. Dal', *Tolkovyj slovar' živogo velikorusskogo jazyka* s.v. *bedrā*; I. Nosovič, *Slovar'*

South Slavic makes it certain that the Proto-Slavic forms must be derived from a base **bedr-*; the variety of combination of common diminutive suffixes does not call for special comment. Miklošič cited *biedrzeniec* s.v. *bedro* 'thigh', without any attempt at justification; and now Holub and Kopečný specifically derive *bedrník* from *bedro*, stating that the root of the plant was an ancient fertility charm because of its goatlike odor, and that the thigh is the seat of reproductive power.¹¹ Today *bedro* is generally accepted, in spite of certain difficulties, as a cognate of Latin *femen*, *femur*.¹² What about the significance of the plant in Slavic beliefs?

The burnet saxifrage was widely used in medieval Germany and in the neighboring Slavic lands as a charm or prophylactic against plague;¹³ in Slavic folk medicine it still serves for stomach ailments, against fever, and generally as a source of strength and good health.¹⁴ German materials note it as a charm for protecting an unborn child and for increasing the flow of milk in a new mother. Less frequent mention is made of the plant as a means of attracting men, and there are only hints that it may have been connected with some sort of fertility cult.¹⁵

Common to all these medicinal and magical uses is the belief that the plant is beneficial, that it will drive away sickness and evil and bring increase and health. This is entirely in accord with the semantic sphere of Skt. *bhadra-* and suggests that both the Sanskrit word and the Slavic names of this plant may be connected with a common IE **bhedr-*.

Passing from plants to insects, we note that the etymological dictionaries give

belorusskogo narečija (St. Petersburg, 1870); Żelechowski, *Ruthenisch-deutsches Wörterbuch* (Lvov, 1886); M. Pleteršnik, *Slovensko-nemški slovar* (Ljubljana, 1894); *Yug. Akad. Rječnik*. Ristić and Kangrga, *Enciklop. srpskohrv.-nemački rečnik* (Belgrade, 1928), give *bédrinac*. Pol. *bedrzyniec* is listed by J. Karłowicz, *Słownik gwar polskich* (Cracow, 1900).

¹¹ F. Miklosich, *Etymologisches Wörterbuch der slavischen Sprachen* (Vienna, 1886); J. Holub and F. Kopečný, *Etymologický slovník jazyka českého* (Prague, 1952). The goatlike odor of pimpernel root is mentioned in *Handwörterbuch des deutschen Aberglaubens* 1.1225, but the thigh as the seat of reproductive power seems to be Holub and Kopečný's own conjecture.

¹² Cf. Berneker; Vasmer; St. Mladenov, *Etimologičen i pravopisen rečnik na bŭlgarskija knižoven ezik* (Sofia, 1941).

¹³ K. Moszyński, *Kultura ludowa słowian 2-1.200* (Cracow, 1934).

¹⁴ Moszyński 220. As an illustration of the survival of the association between *bedrník* and good health, I cite a Czech drinking song which was popular among people of the generation born in the 70's and 80's. It was given to me by Svatava Pirková-Jakobson, and a variant line was supplied by Milada Součková; to both I express my thanks.

<i>Pijme pivo s bobkem</i>	<i>Tak ta bába povídala</i>
<i>jezme bedrník</i>	<i>co do lesa chodívala</i>
<i>nebudeme stonat,</i>	<i>na ten bedrník</i>
<i>(or, budem dlouho živi,)</i>	<i>na ten bedrník.</i>
<i>nebudeme mřít.</i>	

¹⁵ *Handwörterbuch des deutschen Aberglaubens* 1.1223 f. The English pharmaceutical term *pimpernel* root refers to this plant, but the word *pimpernel* has been transferred to the genus *anagallis*. It is interesting that this too has various good-luck and oracular associations; cf. *Standard dictionary of folklore, mythology and legend* 870 (New York, 1950). On the etymology of *pimpinella*, see L. Spitzer in *Word* 7.211-8.

scant attention to various words for the beetle coccinella (usually c. septempunctata) which Americans call the ladybug: Pol. *biedronka*, *biedrunka*, *biedronka*, *biedrzka*; Cz. *bedruňka*; BR *bedrúnka*; Ukr. dial. *bědryk*.¹⁶ Brückner cites only Polish, and associates the words with *biedrzeniec* 'pimpinella' and with his hypothetical adjective **biedry* < *bědr*-. Only Sławski, in his current new work, gives *biedronka* as a separate entry. He follows Lehr-Splawiński, citing Cz. *bedruňka* and interpreting it as an alternate form of *berunka*, *beruška*, diminutives of *beran* 'ram', sometimes applied to the ladybug.¹⁷ There is no mention of any parallel forms in Polish to support this Czech shift, and no explanation of the exceptional change -r- > -dr-. But the aim is to make a recognizable animal of the word, parallel to the SC *božja ovčica* 'god's little sheep', for then the term fits into the astonishing range of names used for this bright-colored little beetle: Eng. *ladybird*, -cow, -fly; Ger. *Gotteskühchen*, *Mariekenperd*, *Herrgottschäfchen*; R *bož'ja koróvka*, Slovene *božji vôlek*, Fr. *bête à [bon] dieu*, and many more.¹⁸

Throughout Europe this insect is well known, and everywhere it has some significance. A list of the headings discussed in the *Handwörterbuch des deutschen Aberglaubens* is sufficient, for our purposes, to show the range of the meanings: from the other world; devoted to the Virgin Mary; bringer of gifts; bringer of children; protection; sunshine; weather oracle and oracle in general; love oracle; bringer of luck; often connected with cow; occasionally used in folk medicine. The connection with divine forces is clearly not only Christian, for it is found also in the Sanskrit term *indragopa* 'protected by Indra', the Yiddish *moise rabeinu's kihele* 'Moses' little cow', and Finnish *ukoinlehmä* 'cow of the god Ukko'.¹⁹ As 'god's animal', the coccinella is a bringer or harbinger of good fortune, often expressed specifically as sunshine and good weather and therefore good crops. Again the proposed IE **bhedr*- fits the form and meaning of the Slavic words, and the developments are regular.

In the whole South Slavic area, the Christian terminology seems to have replaced the names for the ladybug based on **bedr*-. SC *būba mára*, Bg. *božá*

¹⁶ Cf. Karłowicz; Nosovič; Żelechowski; also B. Hrinčenko, *Slovar' ukrains'koi movy* (Berlin, 1924).

¹⁷ F. Sławski, *Słownik etymologiczny języka polskiego* (Cracow, 1952); cf. T. Lehr-Splawiński, *Słownik języka polskiego s.v. biedr[z]onka* (1939). Holub and Kopečný (*Etym. slovník*) list *berunka*, *beruška* as modern forms of an original *bedruňka*, allegedly an 'expressive modification' of *brouček* (itself a diminutive of *brouk*) 'beetle', and give *beroušek*, hypocoristic of *beránek* (dim. of *beran*), as a parallel. The relation of the presumable old root *bręk-/bręk-* to *bedr*- > *ber*- is not further explained. It is a fault of the Holub-Kopečný dictionary that odd forms, which more cautious scholars would frankly label 'unclear', are frequently classed without qualification as 'expressive' forms. Vague semantic relationships are too often preferred to phonetic facts.

¹⁸ *Handwörterbuch d. dt. Aberggl.* 5.1691 ff. gives literature and résumés treating Germany in detail and western Europe in general terms, with some references to other areas. See also *Stand. dict. of folklore* 599, and the excellent and exhaustive monograph *Über die kinderreime vom marienkäfer und dessen benennungen im finnischen und estnischen*, by M. Juvas and K. Vilku, in *Finnish-ugrische Forschungen* 24.154-231.

¹⁹ I wish to thank Morris Halle for calling the Yiddish form to my attention. Ukko was the old Finnish thunder-god and supreme deity; cf. Vilku, *Finn.-ugr. Forsch.* 24.226.

krávica, *litmára*, etc. In the central area of Slavic the old forms are often replaced by diminutives of 'sun', e.g. Cz. *sluníčko*, Sorbian *bože slónčko*, *slyńaško*, Ukr. *sónečko*.²⁰ Both types of substitute show only that the meaning of the word was not vivid enough to express the still living association of the ladybug with divine protection and good weather. Another series of variants can be explained on the basis of an original **bedr-*. Old Czech *vedrunka* is considered by Gebauer and Flajšhans to be the primary form, since it is glossed by *estula*, *estiva*. As the 'summer insect', Flajšhans associates it with Cz. *vedro* 'heat, hot weather'.²¹ There are, to be sure, some cases of *v* > *b* in Czech, but they are exceptional.²² The survival of only the form with *b*, plus the analogues in other languages, argues strongly that it is the change *bedr-* > *vedr-* which needs explanation. Here it is not Czech but general Slavic which gives a more likely guide, for the original meaning of *vedra*, *vedro* is certainly 'clear, fine weather', as it still is in Russian. Precisely this concept is associated with the ladybug: desirable weather, not the unpleasant heat denoted by Cz. *vedro*. This might serve also to explain the loss of *vedrunka* in Czech; the term does not seem apt, and is replaced by completely new names. The Polish substitution of *jedronka* for *biedronka* seems to be simply a distortion which adds no semantic elements and results in a name as arbitrary as the old one.²³

It is interesting to note that in Ukrainian there is, beside the dialectal *bédryk*, also *ščédryk* 'ladybug'.²⁴ In some of the New Year's songs called *ščedrivky* there is an opening formula *ščédryk-védryk*, with the variant *ščédryk-bédryk*.²⁵ Here there seems to be no present-day association with the ladybug, and indeed the second element of the formula is meaningless. The first part is clearly related to *ščedryj* 'generous', and to the widespread Slavic terminology concerning the festivities of Christmas and the New Year. The songs we are concerned with are sung by poor children asking for gifts. The exact relationship between an

²⁰ Cf. Moszyński, *Kult. lud. słow.* 2-1.570; E. Muka, *Słownik dolnosersbskeje rěcy s.v. slyńaško*.

²¹ Jan Gebauer, *Slovník staročeský* (Prague, 1903), lists '*bedrunka* v. *vedrunka*', but the dictionary does not reach the letter *v*. One can only assume that he reasoned like V. Flajšhans, *Klaret a jeho družina* 2.494 (Prague, 1928). It is worth mentioning that Vilks concludes that the primary significance of the ladybug in the Finnish area is as a weather oracle, *Finn.-ugr. Forsch.* 24.228.

²² Gebauer, *Historická mluvnice jazyka českého* 1.423, cites various interchanges of *b* and *v*. To illustrate *b* > *v*, he quotes *vedrník* (14th century) for *bedrník*, without explanation. F. Trávníček, *Hist. mluvn. jaz. česk.* 141 (Prague, 1935), gives no new information for the words in question. It might be noted here that Brückner cites Old Pol. *wiedrzeniec* as an obsolete variant of *biedrzeniec*, offering no explanation for the consonant alternation.

²³ Karłowicz s.v. *biedronka*. *Jedronka* is found in one of the typical verses (in the north Mazovian dialect) asking the ladybug to show, by the direction of her flight, the dwelling of the speaker's future bride; see *Prace filologiczne* 4.823. Other examples of this type of verse in K. J. Erben, *Prostonárodní české písně a říkadla* 89; M. Federowski, *Lud białoruski na Rusi litewskiej* 1.261 (Cracow, 1897).

²⁴ Cf. Żelechowski; Z. Kuzela and J. Rudnyčyj, *Ukrainisch-deutsches Wörterbuch* s.v. *ščedryj* (Leipzig, 1943). George Shevelov tells me that these forms are doubtless from the western Ukraine.

²⁵ See Hrinčenko; Kuzela-Rudnyčyj. I am grateful to George Shevelov for calling my attention to this fact.

older **bedrikъ* meaning 'giver, source of good things' and the ladybug and the midwinter customs is not clear, and deserves special study. But again the base *bedr-* is demonstrated to mean something like 'prosperity, good fortune'.

Polish dialects have a number of names for cows, usually signifying the coloring: *biedrona*, *biedrawa*, *biedrula*, *biedrun*, *biedroń* 'spotted, dappled; red; with a white back'. Slawski cites them s.v. *biedronka* 'ladybug', offering no explicit etymology, while Brückner, noting only *biedroń*, derives it from his hypothetical **biedry*. The terms sometimes seem to denote 'having large haunches'; but since the basic meanings appear to be color, the suggestion that they originally derive from *bedro* 'thigh' is not convincing. Can **bedr-* have some connection with the notion 'cow' so frequently associated with the ladybug? Can it have been perhaps the epithet of a cattle-god, a giver of prosperity? If so, we may suppose that the words derived from it tended, after the loss of the old religious concept, to become confused with derivatives from *bedro*. More information about the sphere of usage and the exact meaning of all these words is necessary before this can be anything but a tentative hypothesis.

At any rate, an IE **bhedr-* seems secure. A few remaining Slavic botanical names can be assigned to the group, even though there is no specific semantic information. Such are SC *bēdrika* 'type of apple', R dial. *bedrjāna* 'a tree (*Cydonia vulgaris*)', Ukr. *bedrjynec* 'peucedanum cervaria', Slovene *bedrīnec* 'a grass; trinia'. As a parallel, one can point to the large number of botanical and zoological terms using *bhadra-* in Sanskrit.

To sum up, OCS *bedrno* seems to be the only attested form of an adjective **bedrnoъ* 'rich', not found in any modern dialect. The base **bedr-* served also for words like **bedrnoъcъ* and **bedrnikъ* meaning 'fortune-bringing plant', applied to the burnet saxifrage, and **bedronka* and **bedrikъ* 'luck-bringing insect', applied to the ladybug, and perhaps for certain other plant and animal names as well. There is no connection with the root *bъd-* 'be awake', and thus this group of words is not evidence for the supposed change of *ъ > ѓ* in Old Church Slavonic or other languages.

NOUN MORPHOLOGY IN ROMANIAN

FREDERICK B. AGARD

Cornell University

0. This paper is a morphophonemic and morphemic study of a class of Romanian words defined by their inflectional behavior as nouns. The point of departure is an array of isolated examples in phonemic form; the result is a classification of nouns by morphological types and a morphemic formula. The corpus from which the examples are drawn consists of Romanian texts¹ interpreted with the aid of native speakers, supplemented with utterances by the speakers themselves.²

The underlying phonemic analysis is the author's. Phonemic data are here given for orientation only; they do not pretend to be a complete presentation of Romanian phonology. The phonemes of standard Romanian³ are the following: SYLLABIC: two front vowels /i e/, three central vowels /i ə a/, and two back vowels /u o/, of which /i i u/ are high, /e ə o/ are mid, /a/ is low; NONSYLLABIC: six labial consonants /p b f v m w/, eight alveolar consonants /t d c s z n l r/, five palatal consonants /č ġ š ž j/, and three velar consonants /k g h/; ACCENTUAL: one stress /' / and three pitches /1 2 3/;⁴ JUNCTURAL: one phrase-internal juncture /+ / and three phrase-final junctures /| || ./.

The language has phonological words, defined as free forms with a single stress, e.g. /kibríturilə/ 'the matches', /deobičéj/ 'usually', /umpriéten/ 'a friend'. It also has morphological words; without recourse to syntax one can isolate four classes by the type of inflection which they exhibit in paradigms: nouns, pronouns, adjectives, and verbs.⁵ We are here concerned exclusively with nouns.

Nouns occur (by definition) in paradigms combining the following morphs: base \pm plural inflection \pm (definite inflection \pm oblique inflection).⁶ Every

¹ W. M. Austin, *Spoken Rumanian* (unpublished): reproduced, by permission of the author and the American Council of Learned Societies, for the School of Language Training, Foreign Service Institute, Department of State; Grigore Nandriș, *Colloquial Rumanian* (London, 1945); Mircea Eliade, *Iphigenia* (Valle Hermoso, Argentina, 1951); Marcel Schönkron, *Rumanian-English and English-Rumanian dictionary* (New York, 1952).

² The author has in preparation a full description of Romanian, based on the same material.

³ By standard Romanian is here meant the dialect of Bucharest as used by the two informants whose speech the author has analyzed: Iris Barbură (B), a native of Arad, Transylvania, who has spent most of her adult years in Bucharest; and Iulian Prundeanu (P), a native of Bucharest. Observation of the speech of numerous other Romanians of various ages, natives of several provinces, has revealed no important structural deviations from the dialect of B and P.

⁴ Since all the forms in this paper are merely cited words, not communicative utterances, the pitch and juncture phonemes need not be included in their notation. The citation contour being uniformly / (1) ... 2' ... 1./, /táta/ stands for /táta¹/ 'father', /kibrít/ for /ki²brít¹/ 'match', and so on.

⁵ There is a residual class of uninflected words.

⁶ The parentheses signify that the definite inflection occurs without the oblique, but not conversely.

noun is a phonological as well as a morphological word; but some phonological words consist of two or more morphological words of which only the last is a noun, e.g. /deobičej/ and /umprieten/ cited above.⁷

Certain limitations on the occurrence of phonemes, statable in terms of phonological words, have a bearing on the morphophonemic part of our analysis. The following phoneme sequences do NOT occur in phonological words: (1) in any position: /əV/, /Və/ except /uə/, /Vea/, /Vi/, /Ve/ except /ie ue/, /éa óa/, /we/, geminate vowels, and /ea/ or /j/ after a palatal consonant (Ç); (2) in word-final: /é ó/, unstressed /i aj/, /Vu/ except /iu/, and clusters with final /l r/; (3) in word-initial: /oa/.

We shall examine the morph classes that appear in nouns in the order of their occurrence, starting with bases.

1. The bare bases are themselves free forms. If we arrange them according to their final phonemes, we arrive at the following groups:⁸

(1) /ə/:			
/cára/ 'land'	/kəməša/ 'shirt'	/pjátrə/ 'stone'	/škoála/ 'school'
/čapə/ 'onion'	/lúnə/ 'month'	/púška/ 'gun'	/táta/ 'father'
/fáta/ 'girl'	/márfə/ 'wares'	/roátə/ 'wheel'	/treába/ 'task'
/fijka/ 'daughter'	/mása/ 'table'	/seára/ 'evening'	/vákə/ 'cow'
/flákərə/ 'flame'	/mína/ 'hand'	/sfíntə/ 'saint'	/várə/ 'summer'
/gíska/ 'goose'	/nóra/ 'daughter-	/slúga/ 'servant'	/viáca/ 'life'
/járna/ 'winter'	in-law'	/sóra/ 'sister'	/wára/ 'time'
/káso/ 'house'	/pisíka/ 'cat'	/stráda/ 'street'	/zəpáda/ 'snow'
(2) /e/:			
/báje/ 'bath'	/hírtie/ 'paper'	/múnte/ 'mountain'	/vréme/ 'time'
/bére/ 'beer'	/kártə/ 'book'	/núme/ 'name'	/wáje/ 'sheep'
/feméje/ 'woman'	/krúče/ 'cross'	/ploáje/ 'rain'	/wáspe/ 'guest'
/fráte/ 'brother'	/mínge/ 'ball'	/soáre/ 'sun'	/žumətáte/ 'half'
(3) /u/:			
/lúkru/ 'thing'	/sédiu/ 'seat'	/sufrágiu/ 'suffrage'	/teátru/ 'theater'
(4) Any permitted final consonant or cluster except /g/:			
/avión/ 'airplane'	/dentíst/ 'dentist'	/kucít/ 'knife'	/pribeág/ 'exile'
/bóc/ 'stick'	/engléz/ 'Englishman'	/kuvínt/ 'word'	/razbój/ 'war'
/bəját/ 'boy'	/ewropeán/ 'Euro-	/már/ 'apple'	/ríw/ 'river'
/bətrín/ 'old	pean'	/muzéw/ 'museum'	/sát/ 'village'
person'	/fíw/ 'son'	/númər/ 'number'	/servíc/ 'service'
/bíč/ 'whip'	/kadów/ 'gift'	/ókj/ 'eye'	/strígət/ 'shout'
/brád/ 'pine'	/kál/ 'horse'	/óm/ 'man'	/studént/ 'student'
/buník/ 'grand-	/káp/ 'head'	/ów/ 'egg'	/tínər/ 'young
father'	/kibrít/ 'match'	/pás/ 'step'	person'
/deál/ 'hill'	/kopí/ 'child'	/pləmín/ 'lung'	/viteáz/ 'hero'
(5) Stressed /i/ or /á/:			
/skí/ 'ski'	/kafé/ 'coffee'	/pižamá/ 'pajama'	
/zí/ 'day'	/pará/ 'money'	/steá/ 'star'	

⁷ These phonological words do not inflect to produce the total paradigm which defines a noun. /deobičej/ is matched only by /deobičejurj/ 'of customs'; /umprieten/ does not inflect at all. On the other hand /običej/ 'custom' and /prieten/ 'friend' are fully inflected.

⁸ All the examples used in the rest of this paper are included here, alphabetically arranged within each group. The accompanying glosses will not be repeated in subsequent citations.

⁹ Informant B says /fíka/, as in her own and P's /pisíka/ 'cat'; but she has /íj/ elsewhere in her speech.

The number of bases with final /ə/ is very large; we need not hesitate to label them ə-bases, and to write them morphophonemically with base vowel -ə-, e.g. *kaśə-*, *lunə-*, *piśtka-*.¹⁰

The number of bases with final /e/ is also very large; we label these e-bases, and write them morphophonemically with base vowel -e-, e.g. *bere-*, *hirtie-kruče-*.

The number of bases with final /u/ is small; in all of them the /u/ is preceded by a cluster which never occurs in final position, or else by unstressed /i/, which likewise is never final. We shall consider these bases, together with the innumerable consonant finals, in sets including also the definite form (meaning 'the ...'), for example /lúkru lúkrul/, /sédiu sédiul/, /muzéw muzéul/, /kadów kadóul/; /án anul/, /sát sátul/, /bíč bíčul/.

We will provisionally assume the definite morph to be /l/ rather than /ul/. If we take the /u/ which consistently appears before the definite /l/ as part of the underlying shape, we have a base vowel for these forms also. Furthermore, where it alternates to final /w/, this is conditioned by the limitation on postvocalic /u/. Its alternation to zero is not phonologically conditioned (cf. /rádu/ a given name, /lupésku/ 'Lupescu'), but it is regular wherever the final consonant or cluster is permitted and is therefore predictable within the subsystem of noun bases. A morphophoneme that includes two alternating members, here /u ~ #/, needs no special notation; we may write these bases morphophonemically with base vowel -u-, e.g. *sédiu-*, *anu-*, *satu-*. We further note that because of the stated limitations, the -u- cannot appear as zero in *lukru-* on the one hand, or in *muzéu-* or *kadóu-* on the other.¹¹

We now have three classes of bases with final vowel: ə-bases, e-bases, and u-bases. The residue of bases ending in /i/ and /á/ is small. For reasons which will appear later, they are better left temporarily unclassified.

2. PLURAL MORPHS. Many ə-bases and u-bases, and a very few e-bases, add a plural morph of the shape /e/ or /je/, for example /kásə káse/, /núme núme/, /sát sáte/, /muzéw muzéje/. Noting that the postvocalic /je/ is conditioned by phonological limitations, we may take the /e/ as basic, and write the plural morph morphophonemically as -e-: *kaśə.e-*, *nume.e-*, *satu.e-*, *muzéu.e-*. The loss of the base vowel is automatic except in u-bases, and in these it is regular.¹²

Many other bases of all three types add plural morphs of the shapes /i ~ j/, /ji ~ j/, or /i ~ #/. In view of such alternations, we must examine paradigms consisting of base, base + plural, and base + plural + definite; for example /lúnə lúnj lúnile/, /bére bérj bérile/, /hirtie hirtj hirtjile/, /sédiu sédi j sédijile/; /krúče krúč krúčile/, /mínge mǐng mǐngile/, /feméje feméj feméjile/. We observe that an /i/ appears consistently in the third form, before the /le/ which

¹⁰ Stress is nonfinal in all these bases, and therefore needs to be written only in those which contain more than one nonfinal nucleus (i.e. vowel, or cluster /ea əs/).

¹¹ It would be possible for the -u- to alternate to final zero in types like *fiu-* or *riu-*; but the pattern is consistently /u ~ w/ after any stressed vowel except in *skiu-* alone, which has the bare base /skí/.

¹² Save in the single instance of /ów ówə/, where, instead of dropping, the base vowel alternates to /w/ and thereby brings about, because of the limitation on /we/, a unique manifestation of the plural -e- as /ə/.

we assume to be the definite morph. This suffix parallels the vowel of the *e*-plural; it occurs as /ji/ under the same conditions, and for the same phonological reason, that the suffix *-e* occurs as /je/. Taking /i/ as basic, we may regard its alternation in final position to /j/ in the first group, and to zero in the second, as conditioned by the limitations on unstressed final /i/ and on /Cj/ respectively. Morphophonemically we write this plural morph as *-i*: *lună.i-*, *bere.i-*, *hîrtie.i-*, *sédiu.i-*; *kruce.i-*, *minje.i-*, *femeje.i-*. The loss of base vowel before *-i* is also automatic, because of the limitation on /Vi/.¹³ The paradigm /serviĉ serviĉij serviĉijile/ fits the pattern if interpreted as *serviĉiu.i-* rather than *serviĉu.i-*; *i*-plurals of this sort correlate with the type *sufraĉiu.i-* rather than with a few *e*-plurals such as *bîĉu.e-*.

There is no way of predicting uniformly which bases add the *e*-plural and which the *i*-plural. In *a*-bases and *u*-bases, neither appears to predominate; but because of the much greater frequency of *-i* in *e*-bases, we class *-i* as the 'standard plural' morph and *-e* as the 'special plural' morph. Bases that take the standard plural are unmarked, e.g. *lună-*, *bere-*, *anu-*; those that take the special plural are marked with the symbol [˘], e.g. *kasă˘-*, *nume˘-*, *satu˘-*.

3. BASE CONSONANT ALTERNATIONS. The base consonant is the consonant or cluster immediately preceding the base vowel. In certain bases of all three classes, the base consonants undergo various alternations, none conditioned but mostly regular within the subsystem. The alternants are in complementary distribution relative to following vowel morphophonemes, i.e. the base vowel or the plural inflection which replaces it.

In *a*-bases and in *u*-bases, there are the alternations /k ~ ĉ/ (except /sk ~ št, šk ~ št/) and /g ~ ĝ/. The non-velar alternants occur before a front vowel, e.g. /vákă vácĉ/, /fijkă fijĉe/, /buník buníĉ/, /gískă gíšte/, /púškă púštj/, /slúgă slúĝ/. Taking the consonant of the bare base as the underlying form, we posit for these alternations the morphophonemes *-k-* (including *-sk-* and *-šk-*) and *-g-*, and write simply *vakă-*, *fijkă˘-*, *buníku-*, *giskă˘-*, *puška-*, *slugă-*. The expected zero alternant of *-i* appears after Ć; this type of alternation cannot occur in *e*-bases, where both base vowel and plural morph are front vowels.

In bases with a standard plural, in all three classes, there are the alternations /t ~ c/ (except /st ~ št/), /d ~ z/, /s ~ š/, and /z ~ ž/. The second alternants occur before *-i*, e.g. /tátă tácj/, /fráte frácj/, /studént studéncj/, /brád brázcj/, /pás páš/, /dentíst dentístj/, /engléz englész/. Again taking the consonant of the bare base as underlying, we posit the morphophonemes *-t-*, (including *-st-*), *-d-*, *-s-*, *-z-*, and write simply *tată-*, *frate-*, *studéntu-*, *bradu-*, *pasu-*, *dentístu-*, *englézu-*. The alternation /s ~ š/ occurs even where *-s-* is the first member of a cluster; as above, the expected zero alternant of *-i* appears after Ć.

In *e*-bases and *u*-bases with a standard plural, there are alternations /l ~ ʎ/ and /n ~ ɲ/. The zeros occur before *-i*, e.g. /kál káj/, /kopîl kopîj/, /plămîn plămîj/. These alternations are not regular; cf. /ĝenerál ĝenerálj/, /án ánj/.

¹³ On the other hand it would not be automatic before a morphophoneme *-j-*, for /Vj/ occurs.

Consequently the morphophonemes here are not simply *-l-* and *-n-*; we symbolize them by *-L-* and *-N-*, writing *kaLu-*, *kopíLu-*, *pləmíNu-*.

Having established the existence of a morphophoneme *-L-*, we now look back at bases whose bare form ends in a stressed vowel, and compare them with their special plurals, e.g. /zí zíte/, /pará parále/, /pižamá pižamále/. The plural forms have a base-alternant with /l/ as base consonant, which can now be identified as the morphophoneme *-L-*; it alternates to zero also before *-ə*. These bases therefore join the *ə*-bases; we write them *ziLə̌-*, *paráLə̌-*, *pižamáLə̌-*. All nouns have now been grouped in three base-classes.

4. AUGMENTED BASES. A few *ə*-bases, at least one *e*-base, and many *u*-bases are augmented by the unstressed sequence /ur/ when pluralized, e.g. /vréme vrémurj/, /lúkru lúkrurj/ /kadów kadóurj/, /kibrít kibríturj/. Since this augment is not regular, we write it into the bases morphophonemically as [ur], e.g. *vrem[ur]e-*, *lukr[ur]u-*, *kibrít[ur]u-*. All bases so augmented have the standard plural.

Six bases are irregularly augmented: *s[ur]orə-* /sórə surórij/, *n[ur]orə-* /nóra nurórij/, *oasp[et]e¹⁴* /wáspe wáspecj/, *kap[et]u¹⁵* /káp kápete/, *oam[en]u-* /óm wámenj/, *mí[j]nə-* /mínə míjnj/.

5. STRESSED-VOWEL ALTERNATIONS IN BASES. In many bases of all three classes there are alternations of the stressed vowel, in complementation relative to the following base vowel or the plural inflection which replaces it. None of these alternations are conditioned, but most of them are regular.

In many *ə*- and *e*-bases with standard plurals, /á/ is raised to /ǎ/ before a high vowel, e.g. /cárə cǎrij/, /kəmásə kəmásǎ/, /márfə márfurj/, /kártə kǎrcj/, /žumətátə žumətǎcj/. On the other hand, /vákə vǎč/, /tátə tǎcj/, /frátə frǎcj/ show that this is not a manifestation of the morphophoneme *-a-*. Since there are no *ə*- or *e*-bases with /ǎ/ as the stressed vowel of the bare base, the alternation may be assigned to a morphophoneme *-ə-*, which regularly appears as /ǎ/ before the high vowel of the plural form but as /á/ before the mid vowel of the base. The bases in question are then morphophonemically *cərə-*, *kəmásə-*, *mǎrf[ur]ə-*, *kǎrtə-*, *žumətǎte-*.

The /ǎ/ of *u*-bases (the only class in which it occurs in the bare base) is regularly fronted to /é/ before a front vowel, e.g. /már mérij/, /bác béce/. Since this alternation is complementary with the alternation /á ~ ǎ/, it can be assigned to the same morphophoneme *-ə-*, and we write *mǎru-*, *bǎcu¹⁵*.

In *ə*-bases and in unaugmented *u*-bases with standard plurals, /eá/ alternates to /é/ before a front vowel or Ć, e.g. /seárə sérij/, /treábə tréburj/, /steá stéle/, /viteáz vitéž/, /príbeág pribéž/. On the other hand, the special-plural *u*-base /teátru teátre/ and the augmented *u*-base /deál deálurj/ do not show this alternation. Since the alternation is regular in terms of base subtypes, it may be

¹⁴ This is merely a free alternant of *oáspete-*, with bare base /wáspete/. The behavior of the morphophonemic cluster *-oa-* is treated below.

¹⁵ The alternation /ə ~ e/ of *-ə-* takes place also in unstressed positions, posttonically in polysyllabic bases, as in *numəru¹⁵* /númər númere/, *tínəru-* /tínər tínerj/. (The morphophoneme *-i-* is explained below.) After a velar consonant /ə/ remains unchanged, as in *flákərə-* /flákərə flákərj/, *strígətu¹⁵* /strígət strígəte/.

assigned to the morphophonemic cluster *-ea-*. We write *searə-*, *treab[ur]ə-*, *steaLə̃-*, *vitedzu-*, *pribedəgu-*, as well as *teatrũ-*, *deal[ur]u-*.

A similar alternation, /á ~ é/, appears under the same conditions and in partial complementation with /eá ~ é/ relative to the immediately preceding phoneme: it occurs after Ç or /i/, but /eá ~ é/ does not; e.g. /čápə čépe/, /járnə jérnj/, /pjátrə pjétre/, /bəját bəjécj/, /viácə viécj/. Since /Čea/ does not occur, /á/ can be regarded as a conditioned alternant of basic /eá/, and the alternation /á ~ é/ after Ç can be assigned to the same morphophonemic cluster *-ea-*. We write *čeapə̃-*, *jearnə-*, *pjeatrə̃-*, *bəjedtu-*, *viedcă-*.

However, both /eá ~ é/ and /á ~ é/ occur after a labial consonant and are therefore in contrast there, e.g. /ewropeán ewropénj/, /kaféá kaféle/ as against /zəpádə zəpézj/, /fátə féte/, /várə vérj/, /másə mése/. If we write *ewropednu-*, *kafedLə̃-*, then the contrasting cases of /á ~ é/ after labials cannot be morphophonemically the same: either the stressed vowel is different, or the preceding consonants are different. Arbitrarily, we regard the preceding consonants as morphophonemically palatalized labials, after which *-ea-* appears as /á ~ é/. We write *zəp'eddə-*, *f'eatə̃-*, *v'earə-*, *m'easə̃-*.

In ə- and e-bases with standard plurals, /oá/ alternates to /ó/ before -i-, e.g. /škoálə škólj/, /roátə rócj/, /soáre sórij/, /ploáje plój/. The alternation is regular, hence predictable in terms of a morphophonemic cluster *-oa-*. We write simply *škoalə-*, *roatə-*, *soare-*, *ploaje-*. A similar alternation, /wá ~ ó/, appears under the same conditions and in complementation with /oá ~ ó/ relative to what precedes: /oá ~ ó/ occurs only after a consonant, /wá ~ ó/ only word-initially, as in /wárə órj/, /wáje ój/. If we take /oá/ as basic, the occurrence of /wá/ is conditioned by the limitation on initial /oa/. We write *oarə-*, *oaje-*.¹⁶

In u-bases with special plurals only, /ó/ regularly alternates to /oá/ before -e-, e.g. /avión avioáne/, /rəzbój rəzboáje/. This alternation is in complementation (relative to bases) with that described above. It can be assigned to the same morphophonemic cluster *-oa-*, which then appears as /ó/ before a high vowel and as /oá/ elsewhere. We write *aviodnũ-*, *rəzboájũ-*.

In a few ə-bases and u-bases, /í/ alternates to /i/ before a front vowel, e.g. /sfíntə sfínte/, /tínər tínerj/, /kuvínt kuvínte/. This alternation is not regular; cf. /bətrín bətrínj/ and /kucíft kucíte/. It is therefore assigned to a morphophoneme other than -i- or -i-. We represent it with the symbol -i- and write *sfíntə̃-*, *tínər-u-*, *kuvíntũ-*.

6. DEFINITE MORPHS AND OBLIQUE MORPHS. The definite inflection has eight different phonemic shapes, all in complementation relative to base types, or to the presence or absence of a preceding plural morph or a following oblique morph, or to both. The oblique case morph is *-or-* /or/ in the presence of a plural morph, otherwise *-j-* /j/. The alternations of the definite are as follows:

(1) /a ~ i/ in ə-bases with standard plurals and in many e-bases, relative to the presence or absence of the oblique, e.g. *vakə-* /váka váčij/,¹⁷ *lunə-* /lúna

¹⁶ Cf. also *oasp[et]e-*, *oam[en]u-*, cited in §4.

¹⁷ /váčij/ means, in isolation, 'to the cow'.

lúnij/, *strada*- /stráda strázij/, *treab[ur]ə*- /treába trébij/, *skoala* /škoála škólij/, *bere*- /bérea bérj/, *kručə*- /krúča krúčij/, *baje*- /bája bájjj/.

(2) /a ~ e/ in ə-bases with special plurals, and in e-bases with /i/ preceding the base vowel, relative to the presence or absence of the oblique, e.g. *kasə*[✓]- /kása kásej/, *kafədLə*[✓]- /kaféáwa kafélej/,¹⁸ *hirtte*- /hirtía hirtíej/. In types (1) and (2) the constant /a/, together with the limitation on final unstressed /aj/, point to the /a/ as basic. We note that a base vowel ə is phonologically lost, as is e after Ç or /i/ and before /i/, e.g. /bérea bérj/. Note further that in the ə-bases and e-bases treated in (1) and (2), the base morphophonemes -k- (-sk-), -g-, -t- (-st-), -d-, -s-, -z-, as well as the stressed vowel morphophonemes -ə-, -ea-, -oa-, -i-, display the same alternations as under the phonemically identical conditions of base + plural described in §3 and §5.

(3) /le ~ lu/ in a number of e-bases, relative to the presence or absence of the oblique, e.g. *frate*- /frátele fráteluj/, *munte*- /múntele múnteluj/, *nume*[✓]- /númele númeluj/.

(4) /l ~ lu/ in all u-bases, relative to the presence or absence of the oblique, e.g. *anu*- /ánul ánuluj/, *deal[ur]u*- /deálul deáluluj/, *məru*- /márul máruluj/, *satu*[✓]- /sátul sátuluj/, *teatru*[✓]- /teátrul teátruluj/. In types (3) and (4) the constant /lu/, as well as the behavior of -u- as a base vowel, points to /lu/ as basic.

(5) /le ~ l/ in ə-bases, many e-bases, and u-bases with special plurals or *ur*-augment or with final unstressed /iu/: all in the presence of the plural, relative to the presence or absence of the oblique, e.g. *lunə*- /lúnile lúnilor/, *kasə*[✓]- /kásele káselor/, *bere*- /bérole bérilor/, *hirtte*- /hirtíjile hirtíjilor/, *satu*[✓]- /sátele sátelor/, *lukr[ur]u*- /lúkrurile lúkrurilor/, *sédiu*- /sédijile sédijilor/.

(6) /j ~ l/ in the e-bases mentioned in (3), and in u-bases other than those mentioned in (5)—all in the presence of the plural, relative to the presence or absence of the oblique, e.g. *frate*- /fráci j frácilor/, *munte*- /múncij múncilor/, *okju*- /ókij j ókijilor/, *məru*- /méri j mériilor/. We might take /le/ as basic in type (5) and /j/ as basic in (6), or /l/ as basic in both. Provisionally assuming separately basic /le/ and /j/, we have a set of four complementary underlying morphs, /a lu le j/. The fact that /l/ precedes a vowel in both of the two-phoneme morphs points to the morphophoneme -L-, which here is zero before /a/, as elsewhere before /ə/ and /i/. Accordingly, we morphophonemize the alternation as -La- ~ -Lu- ~ -Le- ~ -Li-.¹⁹

7. DISCONTINUITY OF BASE AND PLURAL MORPHS. The four morphophonemic shapes of the definite inflection, as just set up, contain vowels which are fully determined by their environment (the base-type and the presence or absence of the plural). These vowels are not in contrast with each other, and hence need

¹⁸ Regarding the /w/ of /kaféáwa/, see fn. 19.

¹⁹ The /w/ which occurs in the definite forms /kaféáwa/, /steáwa/, etc. can be classed as a unique manifestation of the base consonant -L-, appearing instead of the expected zero before /a/ and providing a contrast between the bare base /kaféá/ and the definite /kaféá+a/. Because of the limitation on geminate vowels, there can be no */kaféáa/. The unique *ziLə*[✓]- could, phonologically, have the definite form */zifa/, but it patterns with the other members of its base type (subtype) in having /zifwa/.

not be counted as morphemically relevant. We therefore posit that the definite inflection has merely the morphophonemic shape *-L-*, and that the vowels which follow this *-L-* are SHARDS of a DISCONTINUOUS morph: *-a-* and *-u-* are shards of the base, while *-e-* and *-i-* are shards of the plural inflection. The morphophonemic statements below are based on the assumption of discontinuity.

Bases have shards as follows when the definite *-L-* is present:

(1) *a*-bases have $\cup a$ -, appearing as final /a/ \sim pre-oblique /i \sim e/, e.g. *lună.L.a(j)* = /lúna lúnij/; *kasă.L.a(j)* = /kása kásej/; *steaLă.L.a(j)* = /steáwa stélej/. These are standard-shard bases and need no special morphophonemic notation.²⁰

(2) Most *e*-bases have $\cup a$ -, appearing as in (1), e.g. *bere.L.a(j)* = /bérea bérjij/; *kárte.L.a(j)* = /kártēa kórcij/; *hîrtie.L.a(j)* = /hîrtía hîrtíej/. These also are standard-shard bases.

(3) Some *e*-bases have $\cup u$ -, appearing as final /e/ \sim pre-oblique /u/, e.g. *frate.L.u(j)* = /frátele fráteluj/; *nume.L.u(j)* = /númele númeluj/. These are special-shard bases, which we write by including the shard with the base, thus *frate $\cup u$ -*, *nume $\cup u$ -*.

(4) *u*-bases have $\cup u$ -, appearing as final zero \sim pre-oblique /u/, e.g. *anu.L.u(j)* = /ánul ánuluj/; *satu.L.u(j)* = /sátul sátuluj/; *lúkr[ur]u.L.u(j)* = /lúkrul lúkruluj/. These are standard-shard bases.

The plural morphs have no shards when the definite inflection *-L-*, appearing as /l/, is followed by the oblique *-or-*, e.g. *kasă.e.L.or* = /káselor/, *anu.i.L.or* = /ánilor/.²¹ But when the oblique inflection is not present, the shards are as follows:

(1) The *i*-plural has $\cup e$ -, appearing as /e/, after a standard-shard *a*-base, e.g. *lună.i.L.e* = /lúnile/, *băje.i.L.e* = /bójile/; after any *ur*-augmented base, e.g. *treab[ur]ă.i.L.e* = /tréburile/, *deal[ur]u.i.L.e* = /deálurile/; and after an *u*-base with final unstressed /iu/, e.g. *sădiu.i.L.e* = /sédijile/, *serviciu.i.L.e* = /servíciijile/.

(2) The *i*-plural has $\cup i$ -, appearing as /j/, under conditions other than those of (1), i.e. after a special-shard *e*-base or an *u*-base with a standard plural (except one that has final unstressed /iu/),²² e.g. *frate.i.L.i* = /fráciij/, *anu.i.L.i* = /ánij/, *koptLu.i.L.i* = /kopíijij/, *fiu.i.L.i* = /fíijij/.

(3) The *e*-plural has $\cup e$ -, appearing as /e/, e.g. *kasă.e.L.e* = /kásele/, *nume.e.L.e* = /númele/, *satu.e.L.e* = /sátele/.

8. FINAL CLASSIFICATION. The nouns cited in §1 are classified as follows:

(1) *a*-bases.

Standard: *cără-*, *flăcăre-*, *jearnă-*, *kămășă-*, *lună-*, *oară-*, *pistcă-*, *pușcă-*, *roată-*, *seară-*, *slugă-*, *stradă-*, *șkoală-*, *tată-*, *vacă-*, *v'eară-*, *viedcă-*, *zăp'edă-*.

²⁰ One *a*-base, namely *tată-*, has either the standard shard $\cup a$ - or the special shard $\cup u$ -, appearing as in (4): *tată.L.a.(j)* = /táta tácij/ or *tată.L.u.(j)* = /tátel táteluj/.

²¹ If shards were present, we would not expect them to have zero alternants before oblique *-or-*, for both /eo/ and /io/ occur.

²² Also in the unique case of *tată-* (or *tată $\cup u$ -*, cf. fn. 20): /táciij/.

Special plural: *čepǎ*˘-, *f'eatǎ*˘-, *fijkǎ*˘-, *giskǎ*˘-, *kafedLǎ*˘-, *kasǎ*˘-,
m'easǎ˘-, *parǎLǎ*˘-, *pižamǎLǎ*˘-, *pjeatrǎ*˘-, *sfintǎ*˘-, *steaLǎ*˘-, *ziLǎ*˘-.

Augmented: *mǎrf[ur]ǎ*-, *mǎ[j]nǎ*-, *n[ur]orǎ*-, *s[ur]orǎ*-, *treab[ur]ǎ*-.
 (2) *e*-bases.

Standard: *bǎje*-, *berē*-, *femēje*-, *hǎrtle*-, *kǎrte*-, *kruče*-, *minče*-, *oaje*-, *ploaje*-,
žumǎtǎde-.
 Special shard: *frate*~*u*-, *munte*~*u*-, *soare*~*u*-.
 Special shard, special plural: *nume*˘~*u*-.
 Special shard, augmented: *oasp[et]e*~*u*-.
 Augmented: *vrem[ur]e*-.
 (3) *u*-bases.

Standard: *bǎjedǎtu*-, *bǎtrǎnu*-, *bradu*-, *buntku*-, *dentǎstu*-, *englǎzu*-, *evropeǎnu*-,
fiu-, *kaLu*-, *koptLu*-, *mǎru*-, *okju*-, *pasu*-, *plǎmǎNu*-, *pribeǎgu*-, *sǎdiu*-,
servǎciu-, *skiu*-, *studǎntu*-, *sufrǎgiu*-, *tinǎru*-, *vǎledzu*-.
 Special plural: *aviodnu*˘-, *bǎcu*˘-, *biču*˘-, *kucitu*˘-, *kwǎntu*˘-, *muzǎu*˘-,
numǎru˘-, *ou*˘-, *rǎzboǎju*˘-, *satu*˘-, *strǎgatu*˘-, *teatru*˘-.

Augmented: *deal[ur]u*-, *kadǎ[ur]u*-, *kibrit[ur]u*-, *lukr[ur]u*-, *oam[en]u*-, *rǎ[ur]u*-.
 Special plural, augmented: *kap[et]u*˘-.

9. MORPHEMIC SUMMARY. Bases are morphemically separate from inflections. A base may consist of one or more morphemes—one if it is a root, more than one if it is derived from a root by prefixes or suffixes or both. In either case, a base will be symbolized by *B*.

The plural *-i*- and the plural *-e*- are allomorphs of the same morpheme, symbolized by *I*.

The definite *-L*- is the only allomorph of a morpheme *L*.

The oblique *-j*- and the oblique *-or*- are allomorphs (in complementation relative to a preceding *B* or *I*) of the same morpheme, symbolized by *J*.

In morphemic terms, then, a Romanian noun is $B \pm I \pm (L \pm J)$.

SOME RECENT INTERPRETATIONS OF OLD ENGLISH DIGRAPH SPELLINGS

SHERMAN M. KUHN

University of Michigan

RANDOLPH QUIRK

University of London

1. Three attempts have been made in recent years to apply the techniques of modern descriptive linguistics to the Old English spellings produced by breaking, palatal diphthongization, *i*-umlaut, and velar umlaut. Desirable as such attempts are, the authors of the present article find the results still far from satisfactory.¹

In 1939, Marjorie Daunt proposed a reinterpretation of the OE digraph spellings *ea*, *eo*, *io*, *ie*, claiming that, with a few exceptions, the only real diphthongs of OE were the *ēa*, *ēo*, *īo*, and *īe*, derived from West Germanic *au*, *eu*, and *iu*, and that the conditioned forms of OE (as in *healt*, *heorte*, *liomu*, *ierre*) must be regarded as monophthongal allophones of the *æ*, *e*, and *i* phonemes.² A similar reinterpretation has been incorporated by Fernand Mossé in the phonological section of his recent OE grammar.³ The newest effort in this direction is that of R. P. Stockwell and C. W. Barritt, who limit themselves to maintaining that *æ* and *ea* are merely graphic variants representing one short vocalic nucleus.⁴ We shall try here to ascertain the position which these scholars take, noting some of the implications of their views, and then reexamine the bases of their arguments.⁵

2. All four scholars make it abundantly clear that they find the commonly accepted diphthongal interpretation objectionable, but it is not easy to determine precisely what interpretation they have to offer as a substitute. In the first place, Miss Daunt and Mossé are not clear as to what sounds the digraph spellings represent.

Miss Daunt, at times, argues that the spellings *ea*, *eo*, *io* in conditioned positions represented the simple vowels, which she interprets [æ], [e], [i], with a diacritical symbol qualifying the neighboring consonant (121-2, 128); thus OE *weorpan* is said to have [e] plus a retroflex *r*, West Saxon *feht* [æ] plus a velar fricative, and Anglian *liomu* [i] plus some sort of velarized *m*. At other times, however, she seems to regard such digraphs as representations of diphthongal sounds, 'conditioned diphthongs', in which a short vowel 'acquires a slightly

¹ This article is an expansion of a paper read in the Linguistic Forum of the University of Michigan Summer Session, 23 July 1952.

² Old English sound changes reconsidered in relation to scribal tradition and practice, *Transactions of the Philological Society* 108-37 (1939).

³ *Manuel de l'anglais du Moyen-Âge, I. Vieil-anglais* (Bibliothèque de philologie germanique, Vol. 8; Paris, 1945).

⁴ *Some Old English graphemic-phonemic correspondences—ae, ea, and a* (*Studies in linguistics: Occasional papers*, No. 4; Washington, 1951).

⁵ These scholars are not, of course, the first to question the diphthongal interpretation of OE digraph spellings. Many grammarians assume that in words like *geoc*, the *e* is diacritical; and as early as 1868, W. Scherer suggested that *ea* and *eo* in breaking and velar-umlaut positions indicated no more than that the following consonants were pronounced in a certain way. See *Zur Geschichte der deutschen Sprache* 141 (Berlin, 1868).

diphthongal sound in combination with the off- or on-glide of a preceding or following consonant' (109). She admits that in words like *feoh* (oblique *feo*) the *eo* represents a real diphthong when the *h* is lost (132). She favors interpreting the spelling *ie*, not as a conditioned diphthong, but as [i] (131).

Mossé's indecision is even more obvious. In Old English, he says, *ea*, *eo*, *ie* may represent either diphthongs or simple vowels (30), depending on whether they refer to long or to short nuclei. In speaking of the short nuclei, he states that 'le premier élément vocalique après une consonne palatale et le second devant une consonne vélaire sont des signes diacritiques destinés à indiquer la prononciation palatale, vélaire ou arrondie de la consonne' (31). At the same time, he holds that long or lengthened vowels under these same conditions (as in *gēar*, *scēap*, *nēah*, *mēares*) became diphthongs, phonemically identical with the *ēa* resulting from WGmc. *au* (32, 42). In another place, we find the breaking of short vowels defined as 'la naissance d'une voyelle furtive ou "glide" qui a été notée *a* après *æ* ... *o* après *e* et *i*' (41). Similarly, although we had been led to believe that there was no such sound change as palatal diphthongization, we are informed (43) that this too is a glide phenomenon, the digraphs of which are explicitly contrasted with those in *geoc*, *giong*, *geōmor*. In the latter group 'les graphies *ge-*, *gi-* ne sont pas autre chose que la notation du son [j]', and the sound of the vowel is said to be unaffected.

Stockwell and Barritt take a more definite stand in this matter. They explain the short *ea* by contrasting it with long *ēa* from WGmc. *au*. The latter they describe as '/æ/ plus phonemic off-glide /h/' or phonetically [æh], the off-glide being a part of the syllabic nucleus. The short *ea*, in breaking and velar-umlaut positions, they describe as a 'back allophone' of /æ/ (spelled *e*) plus an off-glide [ə] which they regard, not as a part of the vocalic nucleus, but as part of the following consonant (13). The short *ea* after palatal consonants is also explained as partly vowel, partly consonant, the *e* denoting the palatal value of the consonant, the *a* representing a 'front allophone' of /æ/ (14). We may point out that the use of *e* for a back sound, *a* for a more raised and fronted sound, is not in keeping with the usual practices of Anglo-Saxon scribes. Although the short *ie* (as in WS *giefan*) is outside the scope of their study, this digraph is said to be a representation of the /i/ phoneme, rather than diacritic plus unchanged vowel (26). Thus palatal diphthongization of *e* is treated as both a phonetic and a phonemic change, while palatal diphthongization of *æ* is assumed to be no sound change at all. The analysis by Stockwell and Barritt, upon closer inspection, turns out to be wholly a matter of segmentation: *hēah* they would divide *h-ea-h*;⁶ *neaht*, *n-e-ah-t*; *geat*, *ge-a-t*; and *giefan*, *g-ie-f-a-n*.

3. None of these four faces the problem of the consonantal allophones and allographs which must be postulated if this system of segmentation is correct. Instead of short diphthongs *ea*, *eo*, *io*, and *ie*, we should have a large number of consonant allophones, individually represented by graphs. A description of the consonants of the early West Saxon of King Alfred's time would call for all of the graphs now recognized, plus such additional symbols as *ah* in *meaht*,

⁶ They do not attempt an explanation of *nēah*, in which we have an earlier *æ* in breaking position, or of *gēar*, in which *æ* apparently underwent palatal diphthongization.

al in *healt*, *ar* in *earm*, *ce* in *ceaster*, *ge* in *geat*, *of* in *heofon*, *oh* in *feoht*, *ol* in *seolfor*, *op* in *cleopode*, *or* in *eorl*, *os* in *sweostor*, *sce* in *sceal*. These are by no means all of the new 'consonants' which would be needed; anyone familiar with the Alfredian texts will recall numerous others of less frequent occurrence.

The Anglian texts would require an even greater array of graphic symbols. One illustration will suffice. To represent the consonants in the Vespasian Psalter (British Museum MS Cotton Vespasian A.1.), one would have to set up the following allographs⁷ and interpret them phonetically:

<i>ac</i> in <i>hreacan</i>	<i>oc</i> in <i>spreocan</i>	<i>c</i> in <i>bec</i> and <i>ic</i>
<i>ad</i> <i>feadur</i>	<i>od</i> <i>gebeodu</i>	<i>d</i> <i>bed</i>
<i>af</i> <i>heafuces</i>	<i>of</i> <i>eofur</i>	<i>f</i> <i>cælf</i> and <i>hefig</i>
<i>ag</i> <i>weagas</i>	<i>og</i> <i>weogum</i>	<i>g</i> <i>oferhogan</i> and <i>megen</i>
		<i>h</i> <i>haldan</i> and <i>ðuerh</i>
		<i>i</i> <i>iuguðe</i>
	<i>ol</i> <i>feolu</i>	<i>l</i> <i>hwalas</i> and <i>cild</i>
<i>am</i> <i>freamsum</i>	<i>om</i> <i>liomu</i>	<i>m</i> <i>rum</i>
<i>an</i> <i>deanum</i>	<i>on</i> <i>hionan</i>	<i>n</i> <i>ðonan</i>
<i>ap</i> <i>leappan</i>	<i>op</i> <i>cleopung</i>	<i>p</i> <i>stepum</i>
<i>ar</i> <i>hearpe</i>	<i>or</i> <i>heorut</i>	<i>r</i> <i>hwer</i>
<i>as</i> <i>feasum</i>	<i>os</i> <i>geostran</i>	<i>s</i> <i>blis</i> and <i>ðusend</i>
<i>asc</i> <i>eascan</i>		<i>sc</i> <i>fiscas</i>
<i>at</i> <i>leata</i>	<i>ot</i> <i>uðweotan</i>	<i>t</i> <i>weter</i>
<i>að</i> <i>hreaða</i>	<i>oð</i> <i>cweoðu</i>	<i>ð</i> <i>muð</i> and <i>oðer</i>
		<i>þ</i> <i>þæt</i> and <i>muþe</i>
		<i>u</i> <i>ðuerh</i>
		<i>w</i> <i>sawul</i>
		<i>x</i> <i>oxan</i>

A description of Old English in general would have to include not only the graphs mentioned thus far, but many more which occur in other dialects and texts. It is remarkable that Mossé makes no mention of such graphs, or of the allophones which they would presumably stand for, in his handling of the OE consonants. The other scholars avoid the problem by reason of the limitations placed upon the scope of their studies; nevertheless, we believe that a scholar dealing with a limited problem is responsible for the broader implications of any solution he may propose.

4. On one point the three treatments are internally consistent and in agreement with one another: the graphic distinction between *æ* and *ea* does not reflect a phonemic distinction. Mossé, without offering any clear argument, seems to have been convinced by the fact that *ea* is usually a predictable positional variant of *æ*. In addition, Miss Daunt points to the supposed identity of OE *æ* and *ea* in Middle English. Stockwell and Barritt claim that *ea* and *æ* cannot represent separate phonemes because there are, they say, no minimal pairs so distinguished. These arguments will be dealt with later.

⁷ This list is the irreducible minimum; several other symbols might be needed, depending upon the mode of analysis.

5. The manner in which these scholars treat the whole series of OE sound changes calls for some comment.

Mossé is least aware of the consequences of his statements regarding conditioned diphthongs. He merely repeats the traditional views concerning the relative chronology of breaking, palatal diphthongization, *i*-umlaut, and other changes. He even discusses the first two as vocalic changes, ignoring them in his treatment of the consonants. In short, he presents the traditional sound changes in the traditional order without any attempt to reconcile them with his new theory.

Stockwell and Barritt make an effort to transpose the results of earlier scholarship into their own system. Their principal modification consists in reversing the order of certain changes:⁸ (1) velar umlaut is dated earlier (instead of later) than *i*-umlaut; (2) palatal diphthongization is dated later (instead of earlier) than *i*-umlaut. The first revision ignores strong evidence which has long been familiar to historical linguists. The earliest OE texts show *i*-umlaut virtually, probably altogether, complete; on the other hand, they show fluctuations in spelling which indicate that the influences causing velar umlaut were still active in the language. In the Corpus Glossary (8th-century Mercian), we find *sibun-* (later Mercian *seofen*), *heben-* (later *heofen*), *gelo* (later *geolu*), *quedol* (later *cweodul-*); in the proper names of the Moore MS of Bede's *Ecclesiastical History* (8th-century Northumbrian), we find *herut-* (later Nthb. *heorut*, *hearta*, etc.), *hefen-* (later *heofun*, *heafnas*, etc.). These are but a few typical illustrations; many others could be cited. Moreover, there are several OE words in which a front vowel from *i*-umlaut has undergone velar umlaut; to cite one example, *ondsweorian* in the Vespasian Psalter corresponds regularly to WS *ondswerian* (< **swarjan*). In this instance, the sequence *i*-umlaut—velar umlaut was made possible when the verb shifted from the 7th strong class to the 2d weak class. Further examples may be found in the handbooks.⁹

The second of their revisions, they admit, is unnecessary if *ea* and *ie* are regarded as diphthongs. Their reorganization of the chronology is necessitated by their own interpretation of *ea* as /æ/ and *ie* as /i/, the latter being assumed without evidence (26). Unless they place *i*-umlaut before palatal diphthongization, they are in the awkward position of having to suppose that /æ/ was changed directly to /i/ by *i*-umlaut (whereas elsewhere in Old English, /æ/ was umlauted to /e/): **gæst-* /æ/, by pal. diph. > **geast-* still /æ/, by *i*-uml. > *giest* /i/. They therefore place *i*-umlaut before palatal diphthongization: **gæst-*, by *i*-uml. > **gest-*, by pal. diph. > *giest*. In this rather arbitrary procedure, they ignore the well-known test word for the relative dating of palatal diphthongization and *i*-umlaut. Latin *cāseus* was borrowed early by the English and underwent both changes, resulting in a late WS *cȳse*. The development here must have been *ā* > *ǣ* (fronting) > *ēa* (pal. diph.) > *īe* (*i*-uml.) > *ȳ* (late WS).

⁸ We approve their shifting the second fronting of WGmc. *a* (Luick's 'zweite Aufhellung') to a period after *i*-umlaut. See further R. Girvan, *Angelsaksisch handboek* 79, 81, 83 (Haarlem, 1931); and S. Kuhn, The dialect of the Corpus Glossary, *PMLA* 54.13-6 (1939).

⁹ E.g. K. Luick, *Historische Grammatik der englischen Sprache* 210 (Leipzig, 1921); K. Brunner, *Allenglische Grammatik nach der angelsächsischen Grammatik von Eduard Sievers* 90 (Halle, 1951).

If *i*-umlaut had preceded palatal diphthongization, in accordance with the chronology of Stockwell and Barritt, the development would have been $\bar{a} > \bar{x}$ (fronting)— \bar{x} (*i*-umlaut had no effect upon \bar{x}) $> \bar{e}a$ (pal. diph.), resulting in a form **cēase*.

Miss Daunt goes further than the others, in that she entirely discards breaking, palatal diphthongization, and velar umlaut as sound changes (134). These, she says, 'should be omitted from any account of the sounds of Old English' (122). As a substitute for the traditional series of sound changes, she offers 'progressive palatalization' (134), a general process covering practically all the OE changes in vowels and consonants except nasalization, rounding, and 'loss of sounds and contraction' (136). This substitution of an amorphous theory of development for a set of accurately described sound changes which can be proved or disproved by means of objective evidence seems to us a retreat from that scientific conception of historical linguistics which has produced such revolutionary discoveries since the end of the 18th century. It certainly knocks the props from under many of the criteria for differentiating Old and Middle English dialects.

Unless we have misunderstood Miss Daunt's views regarding the effect of an *i* or *j* on a preceding consonant (109, 129), she has confused two phenomena: the automatic, partial palatalization of a consonant in the neighborhood of *i/j* (as in German *ich* versus *ach*), and the considerably greater palatalization, limited to a specific and datable period, which we know as *i*-umlaut. If, as she seems to suggest, the palatalizing influence was so strong, so immediate, and so progressive that a preceding consonant could not remain velar enough 'to have a distinctly velar on-glide', we must suppose that *i*-umlaut began as soon as any *i* or *j* came to follow an Indo-European stressed vowel.¹⁰

6. From an examination of the theories themselves, we turn now to the bases on which they are founded.¹¹

Both Miss Daunt and Mossé find arguments to justify their interpretations

¹⁰ Yet there is no such universal effect of *i/j* upon preceding consonants, as we may see from modern English *fielding* beside *field*, in both of which many speakers pronounce a velar *l* with a velar or centering on-glide.

¹¹ One point they make is of such a nature that an attempt at direct refutation would lead to no very conclusive or worthwhile results. It assumes that linguists are in possession of universal truths regarding languages in general, and that it is possible to say what can and what cannot exist in language. According to Miss Daunt, 'in existing languages there is, so far as my knowledge and information go, no example of "long" and "short" diphthongs, differentiated only by quantity, being phonemic, and it seems very unlikely that Old English had what is now non-existent' (110). Stockwell and Barritt, after discussing the vocalic nuclei represented by *nama*, *nām*, *nearu*, and *nēar*, say: 'We think it is an accurate statement that no languages have been found that utilize more than three of these four theoretically possible distinctions, or of any other four that might be set up; i.e., three types of syllabic nuclei seem to be a generalizable maximum. On the *a priori* grounds that OE was a real language and ought to act like one, the set of four distinctions is objectionable' (5). We wish merely to offer the following objections. (1) Present linguistic knowledge is probably insufficient for such generalizations. There are literally hundreds of living languages which have not yet been analyzed. (2) This sort of statement seems to deny the possibility that a language may have a unique feature. (3) A-priori reasoning loses its validity when it comes into conflict with objective data. The statements of Miss Daunt and of Stockwell and Barritt regarding Old English must ultimately stand or fall upon the evidence of the MSS.

of the OE digraph spellings by referring to well-known views on consonant color in Old Irish. Mossé treats the matter briefly: 'Il ne faut pas oublier que ce sont des Irlandais qui ont appris aux Anglais à noter par écrit les sons de leur langue. Or en irlandais on se servait également de voyelles diacritiques pour indiquer l'articulation des consonnes voisines' (31). Six years earlier a fuller version of this view had been expressed by Miss Daunt: 'In the first instance, the Irish teachers, listening as foreigners to a strange tongue and trying to write it down, would hear shades of pronunciation which the English speakers would not have heard in themselves, and the Anglo-Saxons would probably try to follow their example and establish an orthographic method and tradition, for native English writing, on their lines' (115).

It would seem that these scholars have started from the accepted fact that the Irish taught the English to write Latin and have gone on to assume that the Irish also taught the English to write English. This assumption, while not inherently improbable, is difficult to verify. It is easy to prove that the Irish were frequent visitors to Anglo-Saxon England, and most of the visitors seem to have been clerics; but there is no early Old English MS in the vernacular which can be traced to an Irish scribe. The first half of the 7th century is the period during which the Irish met the Anglo-Saxons as foreigners and presumably gave them the art of writing. By the close of that century, the English were copying Latin MSS with a skill indicative of long practice in the art. But the earliest MSS wholly in English can be dated no earlier than the 9th century, and even the Anglo-Latin glossaries are no earlier than the 8th. The Irish, also, wrote principally Latin before the 8th century; the earlier Old Irish records consist of a few brief glosses and other fragments.¹² For evidence of Irish influence, one must look to the resemblances between the scribal practices of the two peoples.

Miss Daunt lists six features of OIr. orthography which are paralleled to some extent in Old English. The first five do not concern conditioned diphthongization, and are, as Miss Daunt says, of 'minor importance'. The sixth is the indication of consonant color in Old Irish MSS, a feature which produced forms like *tuaith*, *diglae*, *fiur* (the boldface letters representing the so-called diacritics). From this practice, thinks Miss Daunt, the Anglo-Saxons developed the habit of writing digraph spellings in the neighborhood of certain consonants. Quite apart from any question of similarity between the Old Irish and the Old English phenomena, it should be noted that the English MSS nearest in time and place of origin to the fountainhead of Irish influence show the least development of digraph spellings. Thus spellings like *heben* are common in the proper names of the Moore MS of Bede, spellings like *sparuua* are frequent in the Epinal Glossary, and spellings like *half* are universal in the Corpus Glossary. The spellings which Miss Daunt seeks to attribute to Irish influence, e.g. *heofon*, *spearwa*, *healf*, *fiellan*, do not become regular forms until the 9th century,

¹² On the dates and character of early English and Irish MSS, see E. A. Lowe, *Codices latini antiquiores*, Pt. 2 (Oxford, 1935); E. M. Thompson, *An introduction to Greek and Latin palaeography* 371-402 (Oxford, 1912); W. Stokes and J. Strachan, *Thesaurus palaeo-hibernicus* 1.xiii-xxvi, 2.ix-xl (Cambridge, 1901-3).

the last two in West Saxon, the dialect furthest removed from the great centers of Irish influence. But let us examine the Old Irish spellings more closely.

The most common of the 'diacritics' in Old Irish was apparently *i*, as in *maith*, *teist*, *slóig*.¹³ The influence of this orthographic feature might therefore be expected to appear most strongly in Old English; yet OE forms that might be considered analogous (such as *dæig*) are rare and usually late. The second most common diacritic in OIr. was *u*, as in *daum*, *neurt*, *fiur*.¹⁴ Analogous *u*-forms are extremely rare in Old English; in fact, most texts have none at all. Thirdly there are the Irish *a*-forms, which belong to Middle Irish and in Old Irish are still very rare.¹⁵ Thus spellings like *deacht*, for which it is easy to find OE parallels, were scarcely used by the Irish until after they had appeared in Old English.

Furthermore, the OE digraphs differ fundamentally from such OIr. spellings as *ai* and *iu* in the manner of determining the glide element. In Old Irish, *e* or *a* or *i* can be followed by a *u*-spelling, and *a* or *e* or *o* or *u* can be followed by an *i*-spelling, according to the particular phonetic quality of the following consonant. In OE *ea*, *eo*, etc., the use of *a* or *o* is not determined by the following consonant; instead, *a* appears if the original vowel was *æ*, *o* appears if the original vowel was *e* or *i*. In other words, in Old Irish the choice of the second element in the digraph spelling is determined by the nature of the consonant and not by the original vowel; in Old English it is determined by the original vowel and not by the nature of the consonant.

Finally, we may point out that scholars in the field of Irish are not unanimous in regarding the OIr. spellings as pure diacritics. Thurneysen speaks repeatedly of *i* and *u* as 'glides' (55-7); the case of *i* is elaborated as follows: 'the existence of an audible sound is suggested by the remarkable consistency with which *i* is inserted, and still more by the fact that a word like *immalle*, notwithstanding the palatalized *ll*, is written without *i* (as contrasted with *aill* neut. "other", etc.); in the unstressed pretonic interior syllable full development of the glide could not take place' (56-7).

7. Another argument in support of the new views is drawn from a consideration of Middle English phonology. Miss Daunt and Stockwell and Barritt appear to be under the impression that OE *æ* and *ea* developed identically in Middle English, and that OE *e* and *eo* fell together in a similar manner. 'A supporting reason in favour of regarding the writings *eo* and *ea* as only variants of [e] and [æ], in certain conditions,' says Miss Daunt, 'is that in the overwhelming majority of cases the forms to be traced in Middle English show exactly the development to be expected of [e] and [æ]' (128). Stockwell and Barritt fall into the same error: 'Æ and *ea* fall together in Middle English, resulting in /æ/ in Modern English, when secondary developments do not intervene. When secondary influences operate, they operate to affect *æ* and *ea* in identical ways when all other conditions are identical' (8).

This view derives from some of the old treatments of English phonology, in

¹³ R. Thurneysen, *A grammar of Old Irish* 56 (rev. ed., trans. by D. A. Binchy and O. Bergin; Dublin, 1946).

¹⁴ Ibid. 57.

¹⁵ Loc.cit.

which Old English is represented by the WS dialect of Alfred's time and Middle English by the Southeast Midland of Chaucer's, the two being presented as though the later type of English were directly descended from the earlier. In point of fact, of course, the two dialects belonged to different regions as well as to different periods; Chaucer's *old* is a later form, not of Alfred's *eald*, but of an Anglian *ald*.

A closer examination of Middle English would have revealed to these scholars large bodies of evidence which refute their statements. While the simple short *æ* of Alfred's West Saxon appears as *a* in Southwestern Middle English, there is much place-name material of this area which shows that short *ea* did not develop in the same way.¹⁶ This material has been systematically studied by Hjördis Bohman¹⁷ and Henning Hallqvist,¹⁸ who give many examples of forms which seem to indicate a diphthongal pronunciation; for example, *Estharabyar* < WS *-bearu* (Cornwall, 1324), *Trendelbiare* (Devon, 1314), *Wydebyer* (Devon, 1249), *la Hyele* < WS *healh* (Somerset, 1243), *la Hyales* (Devon, 1280), *Fiernham* < WS *fearn-* (Hants, 14th cent.), *Vialepitte* < WS *fealw-* (Devon, 1316), *Dyalediche* < WS *Dealla-* (Devon, 1281), *Piarrecumbe* < WS *pearroc-* (Devon, 1312). It may be worth while to point out in passing that, if OE *ea* represents a back allophone of /æ/ (as Stockwell and Barritt say), it is strange to find it replaced by ME *ia*, *ya*, *ie*; there are also many forms in which the reduced diphthong is spelled *i* or *y*, indicating that the vowel was high and front. To complete the anomaly, Stockwell and Barritt's front allophone of /æ/ is spelled *a* in Southwestern Middle English, as in *Blakepol* < WS *blæc-* (Devon, 1238).

It is generally recognized that the dialect of the early ME legends of the Katherine Group corresponds most nearly of all Middle English to the Mercian of the Vespasian Psalter. In the gloss to this Psalter, short *ea* represents a breaking diphthong before *r* plus consonant (as in *earm*) and a velar-umlaut diphthong (as in *geðeafung-*); short *æ* appears in several fairly well defined environments in a small number of words, for instance in such smoothed forms as *mæht*; elsewhere WGmc. *a* usually appears as *e*. MS Bodley 34 of the Katherine Group shows the following: (1) *ea* corresponding to the *ea* of the Vespasian Psalter, (2) *a* corresponding generally to the *æ*, and (3) *e* corresponding to the *e*. In the legend of St. Margaret, for example, we find:¹⁹ (1) *earme*, *heardeste*, *hearm*, *feaderes* (gen. sg., VPs. *feadur*), *fearen*, *heatele*, *heatieð*, *beaueieð*; (2) *lahhe* (VPs. **hlæhhan*, cf. *hlæhað*), *mahte*, *mahten*; (3) *dei*, *efter*, *feder* (nom. sg., VPs. *feder*), *hefden*, *schefte*. We do not wish to suggest that these spellings are distinguished

¹⁶ J. E. B. Gover, A. Mawer, and F. M. Stenton, *The place-names of Devon* 1.xxxiii (London, 1931-2). Because of the scarcity of ME literary texts from this area (especially texts which are linguistically reliable), it is necessary to depend largely upon place names for evidence of local pronunciation.

¹⁷ *Studies in the ME dialects of Devon and London* (Göteborg, 1944).

¹⁸ *Studies in Old English fractured ea* (Lund, 1948).

¹⁹ Frances M. Mack, *Seinte Marharete* (EETS, OS 193; London, 1934). Further evidence of the same sort will be found in E. Einenkel, *The life of Saint Katherine* (EETS, OS 80; London, 1884), and in S. T. R. O. d'Ardenne, *Seinte Iulienne* (Paris, 1936). Examples used here have been verified in Ragnar Furuskog, A collation of the Katherine Group (MS Bodley 34), *Studia neophilologica* 19.119-66 (1946-7).

with perfect consistency, although the regularity of the MS Bodley 34 seems extraordinary when one compares it with most MSS of the 13th century. Nor do we propose to interpret the *ea*-spellings of the Katherine Group as diphthongal. Our purpose is merely to show that there is ample evidence that ME scribes did not treat diphthongized and undiphthongized results of WGmc. *a* alike—evidently because they did not sound alike.²⁰

8. As regards the ME developments of *e* and *eo*, there are abundant forms testifying to the fact that ME scribes did not treat them alike. While WS *e* remained in the spelling, the WS short *eo* assumed a number of forms: in addition to *e*, there were *eo*, *u*, *o*, sometimes *ue* and *oe*, these variants being usually interpreted as representing a front-round vowel.²¹ Examples drawn from the place names of Southwest England could be cited at great length, but a few will suffice: *Sturte* < WS *steort* (Dorset, 1327), *Sturtyl* (Dorset, 1250), *Sturte* (Surrey, 1332), *Sturrey* < WS *steorra* (Surrey, 1312), *Chorleton* < WS *ceorl-* (Dorset, 1345), *Churlewod* (Surrey, 1333), *Choerle-* (Devon, 1338), *Cheorle-* (Devon, 1337), *le Hort* < WS *heorot* (Dorset, 1327), *Hurt* (Dorset, 1280), *Hurtescroft* (Surrey, 1336), *Huertemere* (Surrey, circa 1270), *Hurtishole* (Devon, 1299), *Mukeswurde* < WS *meox*, *weorðig* (Devon, 1185), *Hurdewik* < WS *heorde-* (Devon, 1333).²²

In the Mercian area, where in Old English the distinction between *e* and *eo* also obtained, we find a distinction likewise in Middle English. The short *e* remains so spelled, but short *eo* from breaking or velar umlaut of WGmc. *e* or *i* appears in a variety of spellings (*eo*, *o*, *u*, etc.) pointing to a rounded vowel. Again we cite examples from the Bodley MS of the legend of St. Margaret: (1) for breaking of WGmc. *e*: *beornind*, *dorkest*, *eorðliche*, *feor*, *heorte*, *steorren*; (2) for velar umlaut of WGmc. *e*: *beoden* 'prayers', *beoreð*, *eoten*, *heouene* (Royal MS: *houene*), *seoueðe*; (3) for velar umlaut of WGmc. *i*: *cleopieð*, *leome* 'limbs', *neomeð*, *neoðer*, *seonewwen* (or *seonewen*).

In view of such evidence, it is hardly possible to dismiss the OE distinctions *æ/ea*, *e/eo* on the ground that no such distinctions obtained in Middle English. Whatever one's phonetic interpretation of the forms, it is clear that there were ME distinctions parallel to those in Old English, which the scribes recognized and tried to record in their spellings.

9. A third argument for the monophthongal interpretation is based upon a consideration of the OE forms. Miss Daunt and Mossé confine themselves pretty much to the well-known examples given in the grammars; they offer nothing new in the way of evidence, however much their interpretations depart from

²⁰ This discussion, limited to two OE dialects and their ME developments, makes no claim to completeness. An examination of other dialects would reveal further evidence; for example, breaking took place in Kentish as well as in West Saxon, and the Kentish diphthongs may have been preserved as late as the 14th century in such forms as *zyalde* 'sold', *hyealde* 'hold'. See R. Morris, *Ayenbite of Inwyt* (EETS, OS 23; London, 1866).

²¹ See, for example, R. Jordan, *Handbuch der mittellenglischen Grammatik* 86-8 (Heidelberg, 1934); Luick, op.cit. 333.

²² Bohman, op.cit., and B. Sundby, *The dialect and provenance of the Middle English poem The Owl and the Nightingale* (Lund, 1950).

those commonly accepted. Stockwell and Barritt present new material, but in a manner which will probably nullify its effect upon the majority of linguists.

Stockwell and Barritt describe their monograph (3) as a graphemic survey of Old English between A.D. 700 and 900. They give the impression of having examined, and phonemically interpreted, all of the *ae*-, *ea*-, and *a*-spellings in every OE text of the 8th and 9th centuries. They speak of certain forms as 'statistically predictable' (15), which would suggest to us that they have taken fair samplings directly from the MSS or from facsimiles or at least from reliable editions of the texts covered by the study. Yet we are nowhere given a list of the texts or portions of texts which have actually been used, and the two occasions on which they do specify a text show that they have gone outside their prescribed period. Data are cited (15-6) from the gloss to the Lindisfarne Gospels,²³ written in the 10th century, and a single form 'in Bede' is quoted (31), which turns out to be from the Cambridge University Library MS Kk. 3.18, of the 11th century.

They also make statements about early OE phonology which lead one to suspect that their survey is based in part on late texts, WS copies of Anglian originals, and the like. For example, their 'statistically predictable' dialect spellings of six OE words (15) include four in which one can have little confidence. Thus *sceaft* is said to be the typical Anglian form, whereas we have been unable to find any *ea*-forms in Anglian texts of 700-900. Their Anglian *gæt* is extremely rare in the early texts; *ludgæt* in the Corpus Glossary, corresponding to *-gaet* in the Epinal and Erfurt Glossaries, seems to be the only example.²⁴ It is true that one would regard *cæster*, or *caestir*, as typical in an 8th-century Mercian or Northumbrian text, but the 9th-century Vespasian Psalter has only *e*-spellings, and there are no pure 9th-century Northumbrian texts. The form *cæfl* 'halter', said to be the typical WS form, does not appear (to the best of our knowledge) in any WS text of the period 700-900.²⁵

Their mode of describing the practices of the early OE scribes is not likely to reassure historical linguists (including, of course, structural linguists who deal with historical problems). They ignore the fact that the grapheme *x* appears in the MSS in three forms (*x*, *ae*, and *e*), and speak as though the only symbol to be contrasted with *ea* and *a* were *ae* (3). They also suggest that vowel length was indicated in the MSS by a macron. A macron-like sign was used by the scribes, it is true, but primarily as a substitute for *m*, occasionally to indicate an *n* or a contraction. If, for example, a scribe had written the word *la* (i.e. *lā* 'lo') with a

²³ British Museum MS Cotton Nero D.4. Incidentally, their suggestion that previous scholars were unaware of *ea*-spellings after *c*, *g*, *sc* in the Lindisfarne Gospels is unfair to the 'handbooks' which they appear to find deficient. The phenomena in question are dealt with in Brunner-Sievers 63-4; Luick 161; Girvan 68-9; K. Bülbring *Altenglisches Elementarbuch* 62-3 (Heidelberg, 1902).

²⁴ The three glossaries mentioned appear in Corpus Christi College, Cambridge, MS 144; MS Epinal 17; and MS 42 of the Amplonian Library at Erfurt.

²⁵ In words of the type represented by *cæster*, *gæt*, and *sceaft*, 8th-century Anglian texts regularly have *x* (also spelled *ae* and *e*) or *e*; 9th-century Mercian has regularly *e*. The word entered in dictionaries as *cæfl* is attested (in early Mercian and late WS glossaries) in an instrumental form *cæfli* or *caefli*, apparently from **cafuli* and, if so, having secondary *i*-umlaut. An early WS nominative reconstructed from this might be **cæfl* or **ceafl* or **cafol*.

macron over the *a*, his contemporaries would have read an entirely different word, *lām* (i.e. *lām*, 'loam').

10. While presenting old evidence or unreliable new evidence, all four scholars have overlooked many forms which tend to show that the conditioned diphthongs were phonetically distinct from the OE simple vowels. Among the clearest indications is the fact that the short *io* from WGmc. *i* in velar-umlaut position fell together with the short *eo* from WGmc. *e*. The spelling *eo* was commonly substituted for *io* in the 9th century; in later Old English the two diphthongs fell together completely and were regularly spelled *eo*; in Middle English they appear not to have been distinguished.²⁶ Mossé mentions this falling together (46) but fails to see its bearing upon his theory. If the digraph spellings represented unchanged simple vowels, this phenomenon can be explained only as a falling together of simple [e] and [i]. One must then explain why [e] and [i] did not fall together in other positions too.

Both Miss Daunt and Mossé agree that short *ea* and *eo* upon lengthening behaved like *ēa* and *ēo* (< WGmc. *au* and *eu/iu*): *meaŕh*—*mēares*, *feoh*—*fēo*, etc.²⁷ Neither of them satisfactorily explains how lengthening of [æ] and [e] could cause diphthongization in those contexts and not elsewhere.

There are late Old English MSS which contain evidence of another sort; again one example must suffice. The Cambridge University Library MS Ii. 1.33, written soon after 1100, contains part of the Ælfrician translation of Genesis, in which the spelling tradition is less fixed than in some of the better known WS texts. Certain of the forms show consistent departures from the WS spelling tradition and give such indications of phonetic values as to preclude the possibility that *ea* and *eo* were either allographic variants of *æ* and *e* or minute allophonic variants of /æ/ and /e/.²⁸

Broken and unbroken forms of WGmc. *a* are kept clearly apart. The scribe frequently reacted against the *ea*-spelling, not in order to use any symbol suggesting a monophthong, but rather to emphasize the diphthongal character of the sound and to indicate a more palatal first element than that suggested by *e*: *biarn*, *wiaŕð* (also *wierð*, *wiearð*, pret. of *weorðan*), *cyealf*, *getiald* (ppl. of *tellan*), *sialde* (also *syelde*, pret. of *sellan*), *sielt*-, etc. It may be noted that the scribe frequently replaced *ēa* (< WGmc. *au*) with a similar range of spellings: *biagas*, *dieadan*, *gelyafan*, *geliefen*, *lyeum*, etc.

Broken and unbroken forms of WGmc. *e* were likewise kept apart. Occasionally the scribe spelled the broken forms with *u* or *o*: *furen* (for *feorran*), *orþan*. But departures from *eo* are relatively rare; on the contrary, we find the reverse phenomenon. The scribe sometimes used *eo* in place of stressed *o* (as in *geweorhte*,

²⁶ These statements apply to the WS and Mercian dialects; the distinction was retained longer in Northumbrian and Kentish. In Northumbrian, on the other hand, the scribes did not distinguish *ea* and *eo*, although they rather consistently preserved the distinction between *æ* and *e*.

²⁷ Daunt 132; Mossé 31-2, 41.

²⁸ The relevant portions of this MS are accessible in S. J. Crawford, *The Old English version of the Heptateuch* (EETS, OS 160; London, 1922). Crawford studies the significant forms in an appendix to this edition, and in *Trans. Phil. Soc.* 41-7 (1917-20); see also Randolph Quirk, On the problem of morphological suture in Old English, *MLR* 45.1-5 (1950).

weorde) and in place of unstressed vowels (as in *beotwux*, *beo* prep., *þeo* rel. pron.). From these facts we draw some important conclusions. First, it is clear that at the time the MS was written, *eo* in breaking positions represented a sound so far removed from [e] that the symbol could be used for a vowel pronounced with lip-rounding or even for [ə]. Second, the *o* of *eo* was not part of any consonant nor a diacritic of any consonant, but was an integral part of the vocalic symbol *eo*, which could be transferred as a whole from its original predictable contexts.²⁹ Third, and most obviously, broken and unbroken forms of WGmc. *e* had diverged so far as to be recognized by the scribe as being 'different'.

11. Having shown that there is strong evidence in Old English that *ea* and *eo* were phonetically distinct from *æ* and *e*, we are now ready to present some of the evidence which might be taken as indicating that they were phonemically distinct. Those forms cited immediately above, in which *eo* has been transferred from its original limited context to represent other sounds phonemically distinct from /e/, point to a phonemic distinction in the uses of *eo* and *e*.

Other forms could be cited at length from the Vespasian Psalter to show that *æ* and *ea*, *e* and *eo*, were no longer in strict complementary distribution as early as the first part of the 9th century. *Heofen* 'heaven', the usual form of the word, shows the *eo* remaining although the conditioning velar vowel of the final syllable has been weakened to a sound represented by *e*; similarly *seofen*. The plural *steaðelas* has *ea*, either retained from the time when the second syllable had a velar vowel or transferred by analogy from the singular *steaðul*. A similar *ea* which cannot be called conditioned appears in the preterit plural *hneapedum*, beside the singular *hneapade*. The conditioned *ea* in such present-tense forms as *-fearað* is found transferred by analogy to the participle *fearende*. In like manner, we find *beorende*, in which the *eo* is not conditioned, beside *beorað*, *cweoðende* beside *cweoðað*, and *cleopiende* beside preterit forms like *cleopade*. Any phonemic analysis which rejects the distinction between *ea* and *æ*, *eo* and *e*, must account in some way for these non-conditioned forms, as well as many others; the following are selected examples: *amearedes*, *eadesan*, *gedeafenað*, *gleadie*, *gongeweafre*, *hreaðedon*, *spearede*; *gemeodemad*, *geweolegað*, *hehseotle*, *heolstur*, *heoretas*, *ofergeotelas*, *spreocende*.

Although Stockwell and Barritt maintain that there are no minimal pairs in Old English distinguished by *æ* and *ea*, we find a number of pairs which seem to have all the requirements of minimal distinctiveness. Naturally there cannot be many, since short *ea* regularly originated as an allophonic variant of /æ/; but the simplification of certain final consonant clusters, the loss of unstressed medial vowels, and late borrowings from Latin made it possible for minimal pairs to evolve.³⁰ Examples are *ærn* 'house' : *earn* 'eagle'; *bærn* imperative

²⁹ This can be observed sporadically in the codices of OE poetry; e.g. *freom* for *from* 'valiant' (Exodus 14), *weorn* for *worn* (Azarias 185), *weordum* for *wordum* (Phoenix 425).

³⁰ There is a parallel case in the *i*-umlaut of WGmc. *u*, which in a specific context produced a predictable variant of *u*, written *y* and phonetically distinct from the then existing sounds in the OE phonological system. Although minimal pairs developed (e.g. *mund* : *mynd*), they are rare, and it is easier to demonstrate the phonemic identity of *y* by the use of analogous pairs, such as *tyrnan* : *murnan*. On this method of analysis, see

'burn' : *bearn* 'child'; *fær* 'journey' : *fear* 'bull'; (*wudu*)*mær* 'echo' : (*ge*)*mear* 'wicked'; *pæll* 'pallium' : *peall* 'defrutum'; *stæl* 'place' : *steal* 'stall'; *wæl* 'slaughter' : *weal* 'wall'; *wær* 'wary' : *wear* 'a callous'; *wærna* 'wren' : *wearna* gen. pl. 'of hindrances'.³¹

The technique of analogous pairs would uncover further evidence. To show that short diphthongs are distinguished from simple vowels in analogous environments, one could pair off such forms as these: *cwealmes* : *ælmes*, *ears* : *bærs*, *eall* : *pæll*, *hearm* or *bearn* : *hærn*, *bearfendlic* : *hærfest*, *heolfor* : *delfan*, *seolfor* : *selfum*, *heor* : *here*, *meord* or *weorc* : *mersc*, *feorm* : *fersc*. Minimal and analogous pairs of this kind are not necessarily found in any single MS or text, but all are from the same dialect, namely West Saxon, and show no features not found in the MSS of the Alfredian period.

12. In conclusion, we recognize the usefulness of descriptive techniques in the treatment of historical problems, but we believe that the new interpretations of the digraph spellings, as they have been thus far presented, are untenable. Before the new theories can make much headway, certain steps must be taken: (1) the new phonetic values to be assigned to the spellings *ea*, *eo*, *io*, *ie* must be clarified; (2) the consonant phonemes and allophones of Old English must be systematized and brought into harmony with the new vowel system; (3) the whole series of OE sound changes must be explained in terms of the new theories, without recourse to explanations which seem to contradict known facts or which render the changes unverifiable by means of objective evidence.

Even with such clarification and amplification, the theories will remain untenable so long as serious objections to their three major bases remain unanswered. First, the argument from Old Irish rests upon the unverified assumption that the Irish taught the Anglo-Saxons to write English, and upon vague resemblances between two sets of digraph spellings which were fundamentally different. Second, the argument that the OE short diphthongs and short vowels developed identically in Middle English is contradicted by the practice of a number of ME scribes who kept them apart in their spelling. Third, the arguments based upon OE scribal practices either fail to take into account or lightly

B. Bloch, A set of postulates for phonemic analysis, *Lg.* 24.30-2 (1948); id., Studies in colloquial Japanese, IV. Phonemics, *Lg.* 26.86-125 (1950); K. L. Pike, *Phonemics* (Ann Arbor, 1947); id., *Tone languages* 50-4 (Ann Arbor, 1948). As Bloch points out (*Lg.* 26.96), 'Although minimally different pairs are not necessary to prove that a given phonemic difference is distinctive in a particular language, they illustrate such differences more strikingly than other examples.'

³¹ Stockwell and Barritt attempt to explain away some of the pairs made possible by metathesis (31-2). Their first argument is a restatement of the contention that the ME development shows the sounds undifferentiated, a matter already dealt with in this article. Their other arguments are equally weak. They allow a single late spelling *earn* (for *ærn*) to carry as much weight as an abundance of early and irreproachable spellings with *æ*. They argue that *cæse* and *gærs* were so spelled because the scribes wished to distinguish the unpalatalized *c* and *g* from the palatalized varieties; yet the scribes made no such effort in scores of similar cases, such as *cēne* 'bold' and *geŋe* 'going'.

[A reply to this article by Stockwell and Barritt will appear in a later number of the journal.]

dismiss a great deal of evidence bearing upon the phonetic and phonemic interpretation of the digraph spellings. For the present, we see no reason to depart from the established view—namely, that short diphthongs were produced in Old English by breaking, velar umlaut, and palatal diphthongization, or by a combination of any one of these with *i*-umlaut. We have furnished some evidence to show that, within the historical period of Old English, short diphthongs were phonetically distinct from short vowels. We believe that a good case can be made for their having been phonemically distinct as well.

COMPARATIVE NOTES ON YORUBA AND LUCUMÍ

DAVID L. OLMSTED

Yale University

Lucumí¹ is a term here used to refer to a language spoken in Cuba by practitioners of the religion known as *Santería*.² In Cuba both the *santeros* and the language are often called Lucumí, a form of which the origin has been discussed by William Bascom. Bascom seems to have been the first to assert positively that Lucumí is derived from Yoruba, a language of southwestern Nigeria.³ This view is supported by the folklore of Lucumí-speakers, which often contains propositions like these: Lucumí is Yoruba;⁴ Lucumí and Yoruba are the same thing; Lucumí is a dialect of Yoruba; Lucumí is a mixture of Yoruba dialects. Other writers have rejected the classification of Lucumí as related to Yoruba in the usual sense, and have called Lucumí a pidgin language.⁵ The purpose of this paper is to test Bascom's hypothesis by the comparative method—specifically, by investigating phoneme correspondences in cognate items.⁶

1. In the following discussion of the phonemics of Lucumí, [ɛ] represents a class of phones ranging from lower mid to mean mid; [i] represents a class of phones ranging from lower to higher high front; and [u] represents a class of

¹ For acknowledgments see Olmsted, Covert (or zero) morphemes and morphemic juncture, *IJAL* 17.163 fn. 1; id., The phonemes of Yoruba, *Word* 7.245 fn. 1. William R. Bascom encouraged and assisted my field study of Lucumí; funds to support my trip were made available by Mark A. May, administrator of the Ford Foundation grant for research in behavior science to the Institute of Human Relations of Yale University. I am grateful also for the invaluable help of Oswaldo Morales-Patiño, President of the National Archeological Commission of Cuba, of Bernardo Solís, and of Rafael Cepero, and for the cooperation of my nine Lucumí informants, who are unnamed here at their own request. An earlier draft of this paper was read before the meeting of the Linguistic Society of America at Cambridge, December 1952.

² Cf. William R. Bascom, The focus of Cuban Santería, *Southwestern journal of anthropology* 6.64-8 (1950); id., Two forms of Afro-cuban divination, *Acculturation in the Americas* 2.169-79 (*Proceedings of the 29th International Congress of Americanists*; Chicago, 1952); Fernando Ortiz, *Hampa afro-cubana: Los negros esclavos* (La Habana 1916); id., *Los negros brujos* (Madrid, 1917); Melville J. Herskovits, African gods and Catholic saints in New World Negro belief, *American anthropologist* 39.635-43 (1937); id., *The myth of the Negro past* 220 ff. (New York, 1941); Rómulo Lachatañere, *Manual de Santería* (Estudios afro-cubanos; La Habana, 1942); Arthur Ramos, *Las culturas negras en el Nuevo Mundo* 111-39 (Mexico City, 1943); Nicolas Angarica, *El 'Lucumí' al Alcance de Todos* (La Habana, n.d.); Israel Castellanos, *La brujería y el Ñañiguismo en Cuba* (La Habana, 1916).

³ Bascom, The Yoruba in Cuba, *Nigeria* 37.14-20.

⁴ Ibid. 17.

⁵ Cf. Ortiz, *Glosario de afronegrismos* (La Habana 1924); Ramos, *Culturas negras* 137. My findings do not accord with the view, attributed to Castellanos and Ortiz by Ramos (*Culturas negras* 117 fn. 8), that the term *Carabalí* is to be connected with Lucumí or any pidgin language partially derived therefrom. An excellent Carabalí informant, who also called his language *Bricamu* and *Efi*, supplied me with extensive texts. The language is certainly not Lucumí; it may be related to Efik as Lucumí is to Yoruba.

⁶ Cf. Henry M. Hoenigswald, The principal step in comparative grammar, *Lg.* 26.357-64 (1950).

phones ranging from lower to higher high back. Other phonetic symbols, unless otherwise specified, have their usual values. Since in most cases it was not possible to find minimal pairs to prove contrast (nonpredictable occurrence), it is necessary to use environments distributed along a continuum of similarity; in every case, the environments used are the most nearly identical that could be found. The notation /x/y/ is to be read, 'the following forms contain, in the first and second members of each pair, phones tentatively assigned to /x/ and to /y/ respectively, in identical or similar environments'.

1.1. VOWELS. /i/: /i/e/ [jalórdel] 'name of a deity',⁸ [jódi] 'day';⁹ /i/u/ [erí] 'head', [éru] 'black'; /i/a/ [igbó] 'white', [agbó] 'dish'; /i/ε/ [méji] 'two', [méje] 'seven'; /i/o/ [méji] 'two', [méjo] 'eight'.

/e/: /e/i/ cf. /i/; /e/u/ [siré] 'play', [éru] 'black'; /e/o/ [óke] 'hill', [okó] 'man'; /e/ε/ [méji] 'two', [méje] 'seven'; [erí] 'head', [éru] 'black'; [ačelú] 'police', [ejelé] 'bird'; [ačéré] 'maraca', [eré] 'beans'; [éuə] 'rice', [éso] 'fruit'.

/ε/: /ε/i/ cf. /i/; /ε/e/ cf. /e/; /ε/a/ [éči] 'horse', [aja] 'dog'; /ε/o/ [elése] 'foot', [éso] 'fruit'; /ε/o/ [méje] 'seven', [méjo] 'eight'.

/a/: /a/i/ cf. /i/; /a/e/ [ókā] 'one', [óke] 'hill'; /a/ε/ cf. /ε/; /a/o/ [mésa] 'nine', [éso] 'fruit'; /a/o/ [ačá] 'tobacco', [ačó] 'clothing'.

/ɔ/: /ɔ/e/ cf. /ε/; /ɔ/a/ cf. /a/; /ɔ/o/ [ógu] 'hair', [ógū] name of a deity.

/o/: /o/i/ cf. /i/; /o/e/ cf. /ε/; /o/ε/ cf. /ε/; /o/a/ cf. /a/; /o/o/ cf. /o/; /o/u/ [igbó] 'white', [igbu] 'river'.

/u/: /u/i/ cf. /i/; /u/e/ cf. /ε/; /u/o/ cf. /o/.

1.2. CONSONANTS. /b/: /b/gb/ [ibá] 'top (of person)', [igbó] 'white'; /b/f/ [bába] 'father', [fadiséne] 'Sp. tercie'.

/gb/: /gb/b/ cf. /b/; /gb/g/ [agbó] 'dish', [agoló] 'Won't you come in?'; /gb/k/ [arúgbó] 'old', [akukó] 'rooster'.

/kp/: /kp/k/ [akparó] '(?) kind of bird', [ataká] name of a district in Africa.

/k/: /k/kp/ cf. /kp/; /k/gb/ cf. /gb/; /k/g/ [akukó] 'rooster', [bagúra] 'yucca'.

/g/: /g/gb/ cf. /gb/; /g/k/ cf. /k/; /g/kp/ [gagá] 'largest', [akparó] (see above).

/č/, with allophones [s] phrase-initial before [i] and [č] elsewhere: /č/j/ [očū] name of a deity, [ojú] 'eyes'; [ačelú] 'police', [ejelé] 'bird'; /č/t/ [éči] 'horse', [éti] 'ears'; /č/s/ [adáče] 'hand', [anáse] 'birth-cry'; [siré] 'to play', [sf] 'to be, exist'. (My only other example of [s] before [i] is [sinkáfa] 'rice'.)

/j/: /j/č/ cf. /č/; /j/d/ [ojú] 'eyes', [odúa] 'sacrament'.

⁷ [j] and [č] are affricates; [gb] and [kp] are labiovelar stops.

⁸ In view of my promise to the informants not to reveal details of their religion, the Spanish names of Catholic saints corresponding to Lucumí deities are suppressed here. Interested readers can find most of them in the works referred to in fn. 2, particularly Herskovits, African gods.

⁹ These two forms were offered by different informants. I have tried to use only forms common to all informants. Failing that, I have thought it desirable to use pairs of terms, each of which occurred in the idiolect of a single informant. Sometimes not even this was possible.

/t/: /t/ɛ/ cf. /ɛ:/; /t/d/ [méta] 'three', [áda] 'ax'; /t/r/ [méta] 'three', [sarajé] 'clean'.

/d/: /d/j/ cf. /j:/; /d/t/ cf. /t:/; /d/r/ [dadá] 'better', [sarajé] 'clean'.

/r/: /r/t/ cf. /t:/; /r/d/ cf. /d:/; /r/l/ [širé] 'to play', [ilé] 'house'.

/s/: /s/ɛ/ cf. /ɛ:/; /s/f/ [méša] 'nine', [méfa] 'six'.

/f/: /f/b/ cf. /b:/; /f/s/ cf. /s/.

Nasals: /m/n/ñ/ŋ/ [omōmí] 'my children', [egbōniúló] 'they will go', [éñi] 'teeth', [egbōṇiyá] 'aunt'.

The semivowels [i] and [ɥ], characterized by extra-short duration of high front and high back tongue position respectively, and by tongue movement to the position for a following vowel faster than that of [i] and [u], are not found before consonant or pause; the occurrent sequences are [iV ɥV ViV VɥV CɥV]. I have found no contrast between these semivowels and the vowels [i, u]: the latter occur in these positions only with secondary or primary stress, whereas [i] and [ɥ] are never stressed. The semivowels are therefore allophones of /i/ and /u/.¹⁰

1.3. Nasalization is phonemic: [ókā] 'one', [óka] 'bread'. Length is phonemic also: [oró] 'cat', [oró·] 'mango'.

1.4. My material shows three nonpredictably different degrees of stress, correlated with three levels of pitch. High pitch occurs with primary stress, mid pitch with secondary stress, and low pitch with weak stress. The situation is thus different from Spanish, where stress combines (at least in part) with 'segmental' phonemes to form morphemes, while pitch phonemes combine in morphemes only with each other. Lucumí differs equally from Yoruba,¹¹ where pitch phonemes combine in morphemes with 'segmental' phonemes, while stress appears to be nonphonemic. The present analysis takes stress in Lucumí to be phonemic, and pitch to be predictable from stress; the contrary treatment is equally possible. Phonetic high pitch and heavy stress = phonemic primary stress (marked by an acute); mid pitch and medium stress = phonemic secondary stress (marked by a circumflex); low pitch and weak stress = phonemic weak stress (unmarked). Note the following example of contrast between primary and weak: [agbó] 'dish', [ágbo] 'ewe'.

1.5. CLUSTERS. The only initial consonant cluster¹² in my corpus appears in /klíbi/ 'knife'. Every utterance ends with a vowel, but medially I have noted the following clusters: /fr/ in /éfrū/ name of a horse belonging to a particular deity, /nč/ in /pančágara/ 'prostitute', /mb/ in /jembó/ name of a deity, /ŋg/ in /čangó/ id., /rd/ in /jalórde/ id., /ɣm/ in /ā·ɣmalá/ 'flour', /fl/ in /adoflí/ 'lungs', and /ŋk/ in /siŋkáfa/ 'rice'.

Vowel clusters are relatively frequent. Two- and three-vowel clusters are found in all positions; four-vowel clusters occur medially and finally: /euré/

¹⁰ Compare the treatment for Yoruba in Olmsted, Phonemes 248-9.

¹¹ Op.cit. 247.

¹² Since the publication of my article on the phonemes of Yoruba, I have found a few forms with what seem to be consonant clusters in Samuel Crowther, *Vocabulary of the Yoruba language* (London, 1852): *efangelisti* 'evangelist', *alkurani* 'the Koran', *bans* 'marriage bans', *siksik* 'hiccup'. I have not been able to try these out on any informant.

'kid (f.)', /ojoúró/ 'rain water', /agoló/ 'Come in!', /iiá/ 'mother', /mariúó/ 'belt', /màkeuéfô/ 'I write you', /emíkouaúlo/ 'I'm going', /suáiu/ 'onward'; note especially /ouoéuê/ 'paper money'. I have heard no sequence of three or more vowels that does not contain a semivowel allophone of /i/ or /u/.

1.6. The phonemes of Lucumí and Yoruba are then as follows:¹³

LUCUMÍ	i	e	ε	a	o	u	b	gb	kp	d	t	j	č	g	k
YORUBA	i	e	ε	a	o	u	b	gb	kp	d	t	j		g	k
LUCUMÍ	f	s		r	l		m	n	ñ	ŋ					
YORUBA	f	s	š	h	r	l	w	y	m	n	ŋ		1	2	3 4 5 6 7 8 9

2. The sample examined for correspondences contains 231 Lucumí items,¹⁴ most of them probably single morphemes. To choose these, I began with about 200 items from my most reliable informants.¹⁵ On the Yoruba side, I used material recorded from the speech of N. A. Adibi in 1950 and 1951. If a cognate was lacking in this material, I turned to Crowther's Yoruba dictionary. Since this source, though admirable in other respects, suffers from an inaccurate and inconsistent notation of pitch, I had to rely for Yoruba pitch on my Adibi materials. In order to increase the proportion of forms with reliable pitch data, I decided to accept only such Lucumí items as had cognates in the Adibi materials. Of the 231 Lucumí items, 112 proved to have indubitable Yoruba cognates. Twenty more could perhaps be matched with Yoruba cognates on a less rigorous basis, but these were excluded. The proportion of cognates is thus 48.5%.¹⁶

3. The phonemic correspondences found in these 112 pairs of forms are grouped under five headings: (1) certain, (2) probable, (3) based on inconclusive evidence, (4) based on contradictory evidence, and (5) puzzling. The first category includes isomorphisms that occur five times or more; the second includes those that occur two to four times; the third, those that occur only once; the fourth, non-isomorphic correspondences that occur twice or oftener (e.g. cases where a Yoruba phoneme has more than one corresponding phoneme in Lucumí, such as e-e, ε-ε, e-ε, ε-e). The fifth category includes non-isomorphic correspondences that turn up only once.

¹³ The blank spaces in the table are accounted for by the following considerations. Y /š/ = L /č/; Y /h/ and /-/ both = zero (L /-/ does not occur in forms with Yoruba cognates); Y /ŷ/ (two phonemes) = L /ñ/ (a unit phoneme).

¹⁴ Gathered during a field trip to Cuba during June and July, 1952. All informants were from the Habana-Jovellanos-Matanzas area in western Cuba.

¹⁵ The reliability of informants was estimated by checking the internal consistency of their texts, by investigating their language-learning history, and by getting informal ratings of their proficiency in the language from other speakers. This was necessary since the language is learned in adolescence or later, so that age alone is here no guarantee of proficiency.

¹⁶ In marked contrast to the statement which I made in a paper read before the 1952 summer meeting of the Linguistic Society of America at Bloomington. On that occasion, full of zeal after my return from Cuba, I announced that 'there is a surprisingly high percentage of lexical correspondence between Lucumí and Yoruba. ... at a rough guess, about 90%.'

(1)		(2)		(3)		(4)		(5)	
L	Y	L	Y	L	Y	L	Y	L	Y
i	i	ě	š	g	g	ñ	ŷ	e	e
a	a	j	j	i	y	ŋ	ŋ	ε	ε
u	u	u	w	f	f	o	o	u	o
b	b	d	d	n	~r~	o	o	e	i
k	k	s	s			e	ε		
r	r	gb	gb			o	o		
l	l	n	n			ε	e		
m	m	kp	kp			o	o		
						-	-		
						-	-		
						-	-		

Examples of the correspondences:

TRANSLATION	L	Y	CORRESPONDENCES
1. father	bábá	baba	b-b a-a
2. cool	tútú	tutu	t-t u-u
3. cow	malú	ma ₁ lu ₂	m-m l-l
4. is not	kósi	ko si	k-k s-s i-i
5. old	arúgbo	arugbo	r-r gb-gb (o-o)
6. leaf	éué	ewé	u-w (e-e)
7. (a deity)	ěàngó	šango	ě-š (ŋ-ŋ) (g-g)
8. (a goddess)	odúa	odua	d-d
9. eye	ojú	oju ₃	j-j
10. fire	iná	ina	n-n
11. gold, yellow	kpúkpa ¹⁷	kpukpa	kp-kp
12. white	fúfú	fufú	f-f ~-~
13. mother	iiá	i ₂ ya ₂	i-y
14. wife	obini	obiri	n-~r~
15. teeth	eńf	e ₂ ŷi ₃	ñ-ŷ
16. hand	ouó	o ₂ wó ₃	o-o
17. domestic fowl	adié	a ₃ di ₂ e ₂	e-ε
18. speak	sôro	so ₁ ró ₁	o-o
19. ears	eti	e ₂ tí ₃	ε-e
20. fruit	éso	eso	o-o

The first eleven of these pairs exemplify all the correspondences of group 1, and several more besides (shown in parentheses). These correspondences, so far as they obtain between different sounds, can all be accounted for by the principle of regular allophonic change.

4. Correspondences for which the evidence is at least superficially contradictory involve three groups of phonemes: nasalization, stress, and the mid vowels /e ε o ɔ/.

¹⁷ There is also a form /kúkuá/. Some informants have L [kɥ] corresponding to Y /kp/, with [ɥ] lost before [u] but retained elsewhere.

4.1. In fifteen pairs, nasalization in Yoruba corresponds to nasalization in Lucumí. In three pairs a Yoruba oral vowel corresponds to a Lucumí nasalized vowel, and in two pairs a Yoruba nasalized vowel corresponds to a Lucumí oral vowel. Putting the Lucumí representative first, we have: 15 \tilde{V} - \tilde{V} , 3 \tilde{V} -V, 2 V- \tilde{V} . It turns out that the two Lucumí forms with an oral vowel show this vowel in final position, and contain no nasal consonant. The three Lucumí forms in which \tilde{V} corresponds to Yoruba V show this vowel either preceded or followed by a nasal consonant. (These developments did not take place everywhere.¹⁸) Even if we had no ethnohistorical information on Lucumí, these data concerning nasalization would be evidence, as far as it went, that Lucumí resembles Proto-Yoruban less closely than Yoruba does. Assuming that Yoruba is closer to the original, we say simply that Lucumí has lost nasalization in final position when there is no nasal consonant present, and has acquired it in the immediate environment of certain nasal consonants. The opposite assumption would force us to suppose that Yoruba had acquired nasalization in certain randomly distributed forms lacking a nasal consonant, and had lost it in certain nasal environments.

4.2. The correspondences between the nine-phoneme Yoruba pitch system and the three stresses of Lucumí, exemplified by the forty pairs in which there is adequate information for the Yoruba member, are tentatively classified as follows. If we know the Yoruba pitch, we can predict the Lucumí primary stress in terms of a descending hierarchy of rules, each subject to overriding by an earlier rule, and with the general proviso that there is only one primary stress in each item.

The rules: Primary stress in Lucumí corresponds in Yoruba to (1) a rising, falling, or high pitch on a long vowel: L ódara 'very' = Y o₄da₃ra₂; (2) the last high pitch in a form unless this falls on the antepenultimate vowel or earlier: L euré 'goat' = Y e₂wu₃re₃; (3) the final mid in a form containing only mid pitches or a high followed by at least two mids: L ečí 'horse' = Y e₂ši₂; (4) a mid-high rising pitch: L méjo 'eight' = Y e₃jo₂; (5) a mid pitch: L ése 'foot' = Y e₂se₇; (6) a high-low falling pitch: L áčá 'tobacco' = Y a₄ša₇; (7) a low-high rising pitch: L ogú 'San Juan' = Y o₁gū₄. Note how the examples for rules 4 to 7 show the precedence of earlier over later rules.

4.3. Secondary stress follows similar rules: it appears (1) on a rising tone which does not have primary by virtue of the rules above; (2) on a penultimate high tone separated by at least one vowel from the final high; (3) on penultimate low of any form which in Yoruba contains only low tones. Examples: (1) L áčá 'tobacco' = Y a₄ša₇; (2) L ibíbejú 'eyebrows' = Y i₂bí₄bejú₃; (3) L sôró 'to speak' = Y sô₁rô₁.

4.4. The mid vowels /e ε o o/ present the most baffling problem in the comparison of Yoruba and Lucumí. It is to be borne in mind that all speakers of Lucumí are native speakers of Spanish, which has only one mid front and one

¹⁸ Cf. example 10 in the list above, and cf. also /ogú/ name of a deity : /ogū/ the god of iron.

mid back vowel phoneme. The different correspondences observed between the vowels of Lucumí and Yoruba in the mid range, together with the number of items illustrating each correspondence, are shown in the following tabulation:

LUCUMÍ	YORUBA	OCCURRENCES
e	e	10 or more
ɛ	ɛ	10 or more
o	o	10 or more
ɔ	ɔ	9
e	ɛ	9
ɔ	ɔ	6
ɛ	e	5
ɔ	o	1

These figures are based on the whole corpus of 231 pairs. No correlations were observed with any segmental feature of the environment. Of the forty pairs for which reliable information is available on pitch, thirty-four contain mid vowels. Of these, all but six show the same vowel in both members of the pair. One of the six exceptions is perhaps a case of analogical change.¹⁹ Of the remaining five, one pair has L /o/ = Y /ɔ/, where the Yoruba vowel has mid tone and the Lucumí vowel lacks a primary stress; two have L /e/ = Y /ɛ/, where the Yoruba vowel is the only one of its kind in the form and the Lucumí vowel is in final stressed position; and two have L /ɛ/ = Y /e/, where the Yoruba vowel has mid tone and the Lucumí vowel is an unstressed initial. Within this smaller corpus there are no contradictory instances. It is therefore possible that the correspondences among the mid vowels can be explained in terms of the intonational environment.

5. CONCLUSIONS. The data support Bascom's hypothesis that Yoruba and Lucumí are genetically related languages; but there are some unexplained particulars which show that the earlier statement needs revision. First, the correspondences among the mid vowels seem at first glance to indicate a random development of allophones; the problem cannot be settled until we have more evidence on the dialects of Yoruba. Second, the low proportion of indisputable cognates, 48.5%, hardly agrees with Swadesh's hypothesis concerning the rate of morpheme decay, since Lucumí cannot have been separated from Yoruba longer than three hundred years. Third, there is the problem of the origin of Lucumí forms not cognate with Yoruba, of which not one seems to be derived from Spanish. Fourth, the Lucumí vowels and consonants differ but little from the corresponding phonemes in Yoruba cognates—a surprising fact when set beside the apparently rapid lexical turnover.

¹⁹ The form showing analogical change is L /méje/ 'seven', corresponding to Y /eaje/. Contiguous Lucumí numerals are /méta/ 'three', /méri/ 'four', /mánū/ 'five', /méfa/ 'six', /méjo/ 'eight', /mésa/ 'nine', /méua/ 'ten'.

These particulars suggest that the learning-situation in the Lucumí speech community is so atypical as to merit exemption from our usual generalizations about the regular change of allophones. When we have more material from this language to compare with Yoruba, it may turn out that we have oversimplified the possibilities of linguistic culture contact. There may be, in fact, a category intermediate between genetic relation in the usual sense and pidginization; and it may be to this category that we must assign the relation between Yoruba and Lucumí.

REVIEWS

Structure immanente de la langue française. By KNUD TOGEBY. (Travaux du Cercle linguistique de Copenhague, Vol. 6.) Pp. 282. Copenhagen: Nordisk Sprog- og Kulturforlag, 1951.

Reviewed by MURRAY FOWLER, *University of Wisconsin*

This book presents a full-scale examination of a well-known language by the use of a method derived from the postulates of glossematic theory. The principles are those of Hjelmslev (21-2). The method—la méthode immanente—is an analysis of functions designed to produce a non-contradictory, exhaustive, and simple statement. The primary functions¹ recognized are selection, solidarity, and combination; extension is used, particularly in morphology, to establish distinctions. The primary processes are division and classification.

Preliminary to the division and classification is the establishment of the text to be examined (23). Certain elements are excluded: all foreign expressions, since the procedure is considered to be applicable to all languages at once; and all non-phonemic variants. What is required is a text which takes account of all possible commutations. This text, in accordance with glossematic theory, is an infinite one (8).

In the operations of division and classification made upon the text thus established, there are identical problems: in the division, the finding of constituent parts and the statement of their functions; in the classification, the inventory of the elements and the exposition of the system which unites them. The application of the processes in each instance is thus very nearly identical. The members of a functional relationship may, however, appear in one order in the process of division and in the reverse in that of classification.² The end product of this complete examination is the immanent, functional, essential structure of the French language (264).

Within the limits of even a lengthy review it is impossible to give an adequate précis of each part of such a process. In the immediately following paragraphs, therefore, an attempt is made to outline in detail only those portions of the work which appear essential to the understanding of the method. Other portions are

¹ All strict terminology undergoes change when used out of its original context, and no linguistic terminology suffers more in this respect than Hjelmslev's. In this review an attempt has been made either to define or to exemplify terms as they are introduced, in the hope of thus lightening some of the obscurity in which the translation of the French versions into English would otherwise certainly wrap the Danish originals.

Togeby himself has managed admirably to explain his terms as he goes along. For discussion both general and particular of this problem of terminology, see (beside the references in the index to the work under review) André Martinet, *Au sujet des 'Fondements de la théorie linguistique' de Louis Hjelmslev*, *BSL* 42.19-42 (1942-5); Einar Haugen, *Directions in modern linguistics*, *Lg.* 27.211-22 (1951); J. Kurylowicz, *La notion de l'isomorphisme*, *Recherches structurales* = *TCLC* 5.48-60 (1949).

² Syntactically, selection being used as the principle of division, a root presupposes a suffix which defines the formal class of the word; morphologically, selection being used as the principle of classification, a suffix presupposes a root without which it cannot exist (23). See also §3.45 of this review.

also summarized, but in a much more eclectic manner. All criticism is reserved for the conclusion.

The direction of the work is from the total utterance to the smallest constituent elements. There are three divisions in the body of the book: Expression et contenu (25-44); Expression (44-89); Contenu (89-262).

1. Expression and content. The treatment of expression and content considered together as a unit is divided into two parts: syntagmatique (division) and systématique (classification).

The first operation is that of division, and the first problem is the separation of the line of expression from the line of content.³ Between the lines of expression and content the function is solidarity: neither can stand alone. This is the first relationship, and the only relationship, which can be stated at this stage of the analysis (25).

Both in the line of expression and in the line of content that which is accompanied by a unit of intonation is defined as a phrase. Between the phrase and the unit of intonation the function is selection, the phrase presupposing the existence of the intonation (28).

The units of intonation are composed of two elements called 'rising intonation' and 'falling intonation' in the realm of expression, 'incomplete sense' and 'complete sense' in the realm of content. Between these elements the function is selection, the former in each case presupposing the latter (29, 39).

The elements of intonation appear in the inventories both of expression and of content. Since a phoneme of intonation (expression) is always accompanied by a determined morpheme of intonation (content) which it is bound to express and which cannot otherwise be expressed, it is possible to consider them both as a particular kind of elements called morphophonemes (30). Emphases are included within this category (33). Excluded as having neither expression nor content are onomatopoeic expressions (30).

Two preliminary classes of emphasis are recognized: contrastive and emotive. Contrastive emphasis is expressed by the displacement of the peak of intonation normally attached to the penultimate syllable, the new peak of contrastive intonation falling always upon one morpheme no matter whether the initial element be a consonant or a vowel (e.g. *la* ¹*moralité* : *l'* ¹*immoralité*). Emotive emphasis, on the other hand, falls either upon all the syllables of a word or upon the first syllable that commences with a consonant (e.g. *c'est* ¹*in* ¹*sup* ¹*por* ¹*table*; *c'est* ¹*é* ¹*pou* ¹*vantable*) (42). Only contrastive emphasis is considered pertinent to the study of language.

In the first forty pages of the book the establishment of four prime categories of elements is proposed: (1) the phoneme, the element of expression, presupposing the morpheme in the linguistic sign; (2) the morpheme, the element of content, presupposed by the phoneme in the linguistic sign; (3) the morphophoneme, the

³ The reason for this separation is the glossematic one that the units of expression may in some languages not coincide with the units of content: case, number, and gender may be represented together in the same affix (9); or the unit of content may be expressed by a zero sign, as in Latin *cōsul*, where the nominative content and singular content are both represented by zero (32).

element of intonation and of contrastive emphasis, represented both in expression and in content, and therefore both presupposing and presupposed; (4) onomatopoeic sounds, without either content or expression, and therefore with no function of presupposition. In these same pages the morphophonemes are classified in accordance with the function of selection which exists between the elements of rising and falling intonation and between the elements of emphasis and non-emphasis,⁴ the former in each case presupposing the latter.

2. Expression. The treatment of expression is divided into two parts: prosody (division) and phonology (classification).

2.1. The prosodic analysis commences with the definition of the phrase as that which is accompanied by a unit of intonation. The phrase (of expression) is then divided into immediate and mediate constituents: first into groups of syllables, then into syllables, finally into accents and syllabic themes.⁵

The syllable is the smallest unit capable itself of forming a phrase, i.e. capable of being accompanied by a unit of intonation⁶ (44). It has a vocalic nucleus and a consonantal margin. The margin is divisible into initial and final consonants, between which the function is combination, either one being capable of existence without the other (e.g. *le*, *ès*) (55). The vocalic nucleus is composed of a marginal initial vowel (*i*, *y*, *u*), a central vowel, and a marginal final vowel (*i*, *ə*) (65). Diphthongs and long vowels both are analyzed as two vowels united in one syllabic nucleus. The long vowels (following Hjelmslev) are decomposed into the corresponding short vowel plus the *e* called feminine (64), e.g. *jeune* [ʒön] = /ʒø-nə/, *jeûne* [ʒø·n] = /ʒøə-nə/; the elements resulting are in a relationship of combination.

2.2. The inventory of the phonemes follows upon this breaking down of the phrase into its smallest units. The inventory is defined as the smallest number of the smallest elements of which the interchange would produce a change in the content (66). The three principles of division, commutation, and reduction can be deduced from this definition of the inventory (67).

According to the principle of commutation, elements which distinguish two signs otherwise completely identical are phonemes (e.g. *mot* : *pot*) (67). If there are 'also other differences' (67), it is a question only of variants; thus, between the [w] of *trois* and the [u] of *trouas* there is by definition (50) a difference of syllabic frontier; between the [ö] of *jeune* and the [ø] of *jeûne* there is a difference of quantity (67). Division is interpreted in the Hjelmslevian manner: the necessary and sufficient condition of a division is that the parts should be capable of

⁴ 'Non-emphasis' seems to be equated with 'another emphasis' (42).

⁵ All these divisions are not universal. In French there is no function of presupposition between syllables or, since there is no accent (72), between accent and theme (44). The syllable is therefore at once defined as the smallest unit capable by itself of forming a phrase (e.g. *ah!*, *non!*, *moi!*), and the complete process of analysis is mentioned only to show the contrast between French and other languages (44).

⁶ *Pfff!* and similar onomatopoeic expressions, already excluded, are excluded here also by the inability of voiceless consonants to be accompanied by a unit of intonation. Similar vocalic exclamations are, however, included by application of the same criterion and are classed as true interjections.

appearing in an isolated state (68). It follows therefrom that further decomposition into distinctive features is not admissible (70).

2.3. Reduction techniques begin with inventories of initial and final phonemes, the members which appear on the larger of the two lists being tried for identification with the members of the smaller in the following order of procedure: (1) examination of inflectional and derivative forms in which the phonemes may be found alternatively in initial and final syllabic position (e.g. *public* : *publication*); (2) experimental commutation—the placing, for example, of the *-p* of English *tip* before *-in* to see whether the result is *fin*, *sin*, *tin*, or *pin*.

This part of the book concludes with a 'table prosodique' showing the successive decomposition of the text into its constituent parts as far as the division into vowels and consonants.

2.4. The classification (phonology) begins with the division of consonants into vocalic consonants (always appearing next to a vowel) and consonantal consonants (separable from a vowel, for example the *v* in *vrai*). The vocalic consonants are further divided into those compatible and those incompatible with other consonants (i.e. forming or not forming part of a consonant cluster). The consonant *v* can thus be completely defined as an initial consonantal consonant (82). Further classification of consonants is made according to the syncretisms found among them (e.g. *f* and *v*: *f* enters into liaison as *v*). Each final category of consonants will have two members between which there is a relationship of extension: *d*, for example, is intensive, *t* extensive because implied by *d* in liaison⁷ (83). The total number of consonants is 17.

The total number of vowels is ten. These are classified functionally according to their positions within the syllabic nucleus and secondarily according to the syncretisms existing among them.

This part of the book concludes with a 'table phonologique' showing the complete classification of the 27 phonemes.

3. Content. The treatment of content is divided into two parts: syntax (division) and morphology (classification).

3.1. The syntactical analysis commences with the phrase of content, defined by its rapport with a unit of intonation. The inventory to which the analysis is applied is a listing of all the units of the language capable of playing the role of a phrase of content. This inventory is first divided into maximal phrases and other groups capable of being treated as reduced forms of these maximal phrases. The maximal phrase is then divided, by dichotomy if possible, into its immediate constituents, smaller groups being transmitted to a succeeding operation. This process is carried forward until a 'complete'⁸ or typical inventory is made at each stage of the investigation (93).

Togebly emphasizes the importance of the rules of transmission and remission, the formulation of which he credits to glossematics (93).

The number of immediate constituents depends upon their function: if there is a rapport of solidarity or selection, two; if the function is combination, there

⁷ This 'syncretism of liaison' is contrary to glossematic theory (82).

⁸ The quotation marks are Togebly's.

may be more, even to an indefinite number. Commutation can be used only when there is selection or combination; if there is an exocentric relation, the criterion is that of parallelism. In traditional syntactical terms the functions are coordination (combination), subordination (selection), and predication (solidarity) (95).

This procedure provides three ways of establishing a syntactic hierarchy: (1) to have as many levels in the hierarchy as there are operations (cf. the rules of transmission and remission)—in French, fifteen; (2) to take account only of certain functions, so that at each stage at which a certain relationship (e.g. solidarity) can be found there will be a level of the hierarchy; (3) to take account of only one large unit, the phrase, and one small unit, the word, the latter being defined as the smallest unit which can take the place of the former. Each of these possibilities has a certain merit: the enumeration of the operations provides a basis for comparison of languages; the relationship of solidarity is one familiar to traditional grammar; the rapport between the phrase and the word is equivalent to that between the phrase and the syllable on the prosodic (expressional) level (98).

3.2. The inventory of the morphemes follows upon this breaking down of the phrase of content into its smallest units. The inventory is defined as the minimal number of elements minimally commutable (136). As in the case of prosodic analysis, the definition of the inventory of morphemes implies the three principles of commutation, division, and reduction. In practice, however, the difficulties inherent in the discovery of homonyms and synonyms destroy the superficial similarity between phonemic and morphemic analysis. Togeby's solution for the problem of homonyms is to refer first to the principle of division: thus the adverb *pas* contains one morpheme, but the noun *pas* is composed of three: the root, the singular morpheme, and the masculine morpheme (139). He recognizes, however, that strict application of the rule of commutation gives certain unacceptable results which cannot be made more satisfactory by such decomposition. For example, the *-s* of the second person singular in verbs and the *-s* of the plural in nouns are at this state still the same morpheme; and all morphemes having a zero expression remain undifferentiated (140). The functional solution of this crucial problem of morphemic analysis is to choose, between the two possibilities of several morphemes or one morpheme, that which gives the simplest final result.

3.3. The principle of reduction is very briefly treated; but it is clear that Togeby proposes to identify otherwise similar morphemes when they appear in different grammatical positions, such as those of subject and object, and to abstract such elements as the first person in *je* and *me* (142).

This part of the book concludes with a 'table syntaxique' (142) which commences with the text divided again into lines of content and expression and concludes with the group of elements divided into roots and inflectional elements.

3.4. The remaining half of the volume is given over entirely to the classification of the morphemes. The primary bases of this classification are the relations of presupposition found elsewhere. Here, however, in the morphology, these relations are called morpho-syntactic, since they may be characterized equally well either morphologically or syntactically.

Morphologically, a relation may be established between two examples of the same morpheme, for example between the singular morpheme of a noun and of an adjective; this relation is called *homo-elemental* (and also *homo-categorical*). A different relation may be found between two different morphemes belonging to the same category, for example a morpheme of the imperative requiring a morpheme of the subjunctive in a following clause; this relation is *homo-categorical*, *hetero-elemental*. Finally, a relation may exist between morphemes belonging to different categories, such as prepositions governing case-morphemes or [subordinating?] conjunctions requiring certain modal elements; this relation is *hetero-categorical*, *hetero-elemental*.

Syntactically, a relation existing between members of the minimum syntactical unit (defined as root plus inflectional elements) is called *homo-fundamental*. Such a *homo-fundamental* relation cannot be *homo-categorical* because a category is never represented twice in the same group of inflectional elements. Elements belonging to two different categories⁹ may contract the *homo-fundamental* relation called *syncretism*, for example a morpheme of the neuter gender requiring *syncretism* of the nominative and the accusative case morphemes. All other relations are *hetero-fundamental*. A relation within a word is also called in Hjelmslevian terms *homo-syntagmatic*: this relation can never be *homo-categorical*, since the root and the inflectional elements belong by definition to different categories. Relations existing beyond the boundary of the word are *hetero-syntagmatic*. Taking account of the junction¹⁰ as the unit intermediate between the word and the clause, a *homo-junctional* relation is possible; this relation can be *homo-elemental*, as in the agreement of adjective and noun, or *hetero-categorical*, as in the relation between a preposition and the morpheme of case which it requires. A relation within a clause is *homo-nexual*. *Homo-nexual* relations comprise all *homo-junctional* relations, and also such *homo-elemental* relations as that between subject and predicate and such *hetero-categorical* relations as that between verb and case. Finally, there are *hetero-nexual* relations; these may be *homo-elemental*, such as the agreement between antecedent and relative pronoun, and the sequence of tenses in Latin; or *hetero-elemental*, such as the requirement of a subjunctive by an imperative; or *hetero-categorical*, such as the requirement of certain modes by certain verbs, and that of the imperfective aspect by the preterit in French. All these relations are represented in a tabular summary for easy reference (144).

3.41. The morphemes appearing in the inventory are defined first by their *homo-syntagmatic* relations,¹¹ then by their *hetero-syntagmatic* relations in units of increasing extent, the *homo-fundamental* relation being employed only at the end of the procedure (146).

The secondary basis of classification is extension. *Homo-fundamental* extension is called *dominance*, *homo-syntagmatic* extension is *defectivity*, *hetero-*

⁹ Cf. §2.4 and fn. 7 of this review.

¹⁰ Junction is defined in glossematic terms (97, of which this is a paraphrase, not a translation) as the division formed of divisions having intensive nominal inflectional elements, not forming a verbal unit, between which there is a relationship of determination.

¹¹ Glossematics uses the inverse order: first *hetero-syntagmatic*, then *homo-syntagmatic* relations.

syntagmatic extension is direction. The element which combines with the greater number of other elements, or appears in the more varied context, or is represented in the larger number of syntactical types, or (although this is doubtfully stated) has the wider semantic range is extensive; the minor element in each instance is intensive (147-8).

3.42. There are four categories of morphemes: inflectional elements ('flexifs'), roots, derivatives, particles.¹² These are defined by their homo-syntagmatic functions: between the root and the inflectional elements there is a relation of solidarity; derivatives presuppose roots and inflectional elements; particles are neither presupposed nor presupposing (152). Certain problems are immediately solved by reference to these definitions: *-ment, bien, mal*, for example, are regarded as cases (153).

3.43. The inventory of the members of each of the categories begins with the class of inflectional elements. The divisions of this group are the verbal inflections of mode, time, and aspect; the verbo-nominal of number and person; and the nominal of gender, case, and comparison¹³ (158).

The verbal inflections exist in the following hierarchy: mode is independent; tense presupposes the existence of mode, and aspect presupposes that of both mode and tense, though tense is independent of aspect (164). The classification of the modes is as follows: the imperative is intensive, both because it has a defective inflection (196) and because it combines only with person and number, not with aspect and tense; the subjunctive is extensive with respect to the imperative; and the indicative is extensive with respect to the subjunctive (168). By hetero-syntagmatic criteria, the imperative is intensive because it presupposes the subjunctive, and the indicative is neuter because there is no direction between it and other modes. From the reference-point of the context in which they can occur, the subjunctive is intensive because of restrictions on its employment, the indicative extensive. By the criterion of the syntactical types in which they can occur, the imperative is intensive with respect to both the subjunctive and the indicative (171).

The verbo-nominal inflections fall into two groups: person (pronominal and verbal) and number (nominal and verbal) (182). As a result of this classification of person and number, the nominal inflections are restricted to case, comparison, and gender (188).

This part of the book, dealing with inflection elements, concludes with a 'table des flexifs' (201).

3.44. Roots (the second category of morphemes) are defined by the relation of solidarity existing between them and certain inflectional elements (200). Since, however, the relation is one of solidarity, the roots and the inflectional elements are reciprocally defined (202) by a homo-syntagmatic criterion—by their function within the minimal syntactic unit, the word. Thus verbs are defined by their combination with verbal and verbo-nominal inflectional elements, and these

¹² Togeby notes (155) the lack of universality of these categories.

¹³ Declension and conjugation are excluded, for reasons of simplicity, from the inventory of inflectional elements, even though the argument is admitted that they prove their existence by conditioning syncretisms (158).

elements in turn by their relation to verbal roots. For the further classification of verbs themselves, Togeby proposes a hetero-syntagmatic criterion: the capability of being or of not being an auxiliary verb (205). The former of these categories is extensive, since all the verbs in it have the double function of auxiliary and principal verb. The results of the final classification of verbs is tabulated (211-2).

Nouns are defined (214) by their fixed or defective inflection in respect to gender; adjectives by their parallel free or variable inflection (nouns are always masculine or feminine; adjectives are sometimes masculine, sometimes feminine). The property of pronouns (214), but one shared with proper nouns (215), is that of having a zero root. Numerals, placed by Hjelmslev also in the category of zero roots (216), are kept apart by Togeby because there is commutation possible between any two numeral nouns (*deux : trois*). He therefore provides a third criterion, that of 'ordinary' inflection. Nouns thus have a defective inflection in respect to gender, adjectives a parallel free inflection; numerals have a defective inflection in respect to number, nouns a parallel free inflection. The final classification of nouns is into proper nouns, pronouns, numeral nouns, substantives, and adjectives (216-223).

3.45. Derivatives (prefixes, suffixes, and the non-finite parts of the verb) are defined as morphemes which, within the word, presuppose the existence of the root and one or more inflectional elements in a relation of solidarity with the root. These are classified in accordance with their syntactic homo-syntagmatic function. Morphologically, all derivatives presuppose the existence of a root. Syntactically, however, the prefix is subordinated to the root, which has the same function as the entire unit; but the root is subordinated to the suffix, which determines the function of the unit (225). The class of the former alternative (prefixes) is called homogeneous; that of the latter (suffixes) heterogeneous. Homogeneous derivatives are thus presupposing, heterogeneous derivatives presupposed. The non-finite forms of the verb are at once presupposing and presupposed (226) and are called, as the third class of derivatives, homo-heterogeneous.

Tables of prefixes (231) and of suffixes (239) are given.

3.46. The particles remain: adverbs, liaison words, interjections. The inventory contains all uninflected words. The classification scheme adopted up to this point excludes any homo-syntagmatic criterion for the hierarchical arrangement of particles. Togeby here makes use of the terms previously employed in the classification of prefixes and suffixes and arrives at the following schematic arrangement of particles: adverbs are homogeneous particles; liaison words (conjunctions, prepositions) are homo-heterogeneous; interjections are neither homo- nor heterogeneous (245).

To establish the inventory of adverbs, the definition (homogeneous, therefore presupposing or subordinated) is now used to eliminate all particles which do not exist in this relation of presupposition and all those which also have another (e.g. prepositions like *devant* and *derrière*, conjunctions like *quand*) (245). After these and other reductions a total of twenty-six adverbs remains. These are classified first, since they are always subordinated, by reference to the members which

they presuppose; thus there are verbal adverbs, verbo-nominal adverbs (*ainsi, assez, aussi*), and one purely nominal adverb (*très: très malade, il est très artiste*) (247). These are further classified, first as homo- or hetero-nexual, second according to the modal construction which accompanies them, third according to their use with *ne*, and finally according to their extension. As a result of this complex system, the 'table des adverbess' (250-1) is both lengthy and intricate.

Liaison words, which, by definition, have the property of being able to appear in an exocentric construction, are subdivided by Togeby in accordance with their relationship of direction to verbal and nominal inflectional elements.¹⁴ Thus there are subordinating conjunctions (verbal), prepositions (nominal), and coordinating conjunctions (with no such relationship).

There are four subordinating conjunctions, *que, si, quand, comme*, of which the first two are hetero-nexual, the last two homo-hetero-nexual. *Que* can further be differentiated from *si* by the fact that it can be followed by the subjunctive mode, *quand* from *comme* by a different aspectual construction (253-4).

There are twenty-two prepositions. They are classified by the criterion of direction. All are homo-nexual because of their requirement of certain cases; fourteen, however, are homo-hetero-nexual because they may have a completive clause introduced by *que* (e.g. *avant que*). Of these fourteen, six may combine with *que* only through the intermediary of *ce*; not all require the same mode; *avant* and *sans* combine with *ne* in a following clause; *dès* requires that the aspects and the tenses of the principal and the subordinate clauses be identical (256-7). These and other differences serve for a complete tabulation of prepositions (259).

There are seven coordinating conjunctions: *mais, et, ni, ou, or, donc, car*. The last three are always hetero-nexual, as is *mais*, although more uncertainly. These are further divided with reference to their extension, by their relation to *ne*, and by their rapport with the category of number. These relations are shown on the 'table des mots de liaison' (261).

The interjections which now alone remain are distinct by definition from onomatopoeic expressions. They are primarily words of reply; and these, by the elimination elsewhere of all other possibilities, have now been reduced to two: *oui* and *si*. If all expressions like *oh* and *ah*, nearly limitless in number, are now excluded, the analysis of the language ends with the statement that both *oui* and *si* are homo-hetero-nexual, but that *si* is opposed to *oui* by its presupposition of *ne* in a preceding clause.

The results of the entire analysis are represented in a 'table des éléments' (262-3). This and the other tables which appear throughout point up the excellent arrangement of the book. Save for the quotation, scrupulous to excess, of the opinions of others, and for the histories of ideas which encumber every section, the progress of the investigation, thanks to its clear order, can be followed directly from page to page. In a work so compressed as this one, clarity is certainly one of the very highest virtues, and Togeby has evidently worked hard to attain it. There is an excellent index of ideas and definitions (273-8) and a useful table of contents. (There is, however, no index of words.) A brief conclusion

¹⁴ The criterion is borrowed from Holt, who used it to classify the entire category of particles (252 n. 1).

(264-7), itself entitled 'La structure immanente de la langue française', culminates in a suggested typology, on morphological rather than phonological grounds, for all languages. The book ends with a résumé in Danish. Those who cannot afford the really great amount of time necessary for careful reading of the entire work can acquaint themselves with the method and the conclusions by studying the summaries and consulting the indexes and going through the résumé. (The résumé should not be omitted: it is worth the trouble of going through even with grammar and dictionary at hand.) It goes without saying, to be sure, that such a casual survey will not reveal the great consistency which is the essential characteristic of the work.

In an investigation such as this, which begins with a rigid definition of problems not exclusively monolingual, which works with a linguistic method applicable to all languages at once, and which proposes nonetheless to take the 'infinite text' of one language for its province, flaws are inevitable; and some of these flaws, at first sight at least, will appear grotesque. It is remarkable that they are so few in this book.

It is difficult, despite the cogent reasoning, to accept quite readily the interpretation of *loup* as /luf/ (81) and the analysis, however functionally necessary, of a phonetically closed syllable into two open ones and of a phonetically open syllable as a closed one (51). Furthermore, if *nue* is distinguished from *nu* precisely as *petite* is from *petit* (23) and if *cou* and *coup* are homonyms only in pronunciation (139), then it is hard to see how the possibility can be entertained and supported (even though not adopted) that the difference between *sachions* and *sachons* and between *sachiez* and *sachez* is 'de nature purement graphique' (166). The value, moreover, to take an example of a different sort, of such a statement as '*-ier* indique ... un récipient' diminishes to a linguistic zero when it is qualified by '*presque toujours*' (198). And, without definite proof, it is quite impossible to accept the proposal that Latin had neither intonation nor emphasis (31, 266).

It would be unfair to make the foregoing statements without adding at once that the instances given from Togeby's work exemplify certain consequences of his method and hence of his theory, rather than errors in observation or analysis. In the first place, it is still at least doubtful whether the theories of glossematics as modified by Togeby can be substantiated in the inspection of any one language without distorting the ultimate appearance of that language. It is possible, that is to say, to accept tentatively the theoretical, logical, and systematic necessity of interpreting *loup* as /luf/; but it is impossible to believe that the French word *loup* has any equivalent in reality which could be written 'luf'. In the second place, as Martinet has pointed out,¹⁵ the Hjelmslevian school's admittance of the graphic representation of speech sounds to linguistic analysis is perilously close to bringing in the absurd. In the third place, in all other kinds of linguistic analysis, *typical* means representative not of 'complete'¹⁶ but of complete inspection.

As a logical consequence, no doubt, of this kind of application of this kind of method, there is an air of unreality about the whole work. Without impugning

¹⁵ In the article cited in fn. 1 of this review.

¹⁶ The quotation marks are Togeby's.

the author's integrity or even sincerity of purpose, it is still possible to wonder whether the analysis of French can really begin with the breaking of an 'infinite text' into modulation groups. In a unidirectional process such as this, a start must be made at a place where nothing whatever is known. One would like to see such a start really made: to see a sizeable unit of French—say half a million phones—broken into groups by a linguist genuinely ignorant of the language, solely by reference to the intonation, and to observe the sequences thus furnished subjected to the next step in the analysis without change and without diminution. It is possible likewise to question whether Togeby has solved the problem of the morpheme by reference to the criterion of simplicity (140); as for the 'domaine d'emploi', it can be used only to define an object already isolated by other means: that is a simple principle of distributional analysis. A similar question may, indeed, be asked concerning the phonemic method, since it is certainly not made clear how anything at all is revealed by placing the *-p* of the English word *tip* before *-in* to see whether the result is *fin*, *tin*, *sin*, or *pin*, unless either the answer to the problem is known beforehand (in which case this is only a kind of game), or unless a rigorous phonetic analysis at least accompanies the experimental commutation (and in that case, of course, it is not linguistics without phonetics).

In the end, therefore, the question which this review must raise is whether Togeby's work reveals the actual structure of the French language or demonstrates, instead, the possibility of imposing upon French a structure previously conceived. Without preparing for comparison another analysis as detailed, as methodical, and as consistent as Togeby's, it is impossible to give an answer. It may be surmised, however, that the enduring value of work such as this will be found to arise less from care for the ultimate detail than from constant attention to the postulated concept of the whole; and it may be, therefore, that this is the first full-scale demonstration of a new way of knowing things well-known, but by a different way, before. But that is a problem of epistemology which the science of linguistics has thus far not completely solved. For the present it must suffice to pay due tribute to the undoubted brilliance of Togeby's work, and to the two virtues of consistency and clarity which shine more brightly through it at each successive reading.

Die innere Form des Deutschen: Eine neue deutsche Grammatik. By HANS GLINZ. (Bibliotheca germanica, Vol. 4.) Pp. 504. Bern: A. Francke AG Verlag, 1952.

Reviewed by WILLIAM G. MOULTON, *Cornell University*

The traditional analysis of German grammar that is used in the schools of Germany, Austria, and Switzerland is codified in a work known as the Duden grammar, after its former editor, Konrad Duden. In origin, the analysis goes back over a hundred years to one Karl Ferdinand Becker, who in 1831 published a *Schulgrammatik der deutschen Sprache* that was based fundamentally on logic. Becker's analysis was taken over into the eighteen editions of Friedrich Bauer's *Grundzüge der neuhochdeutschen Grammatik für die unteren und mittleren Klassen*

höherer Bildungsanstalten (1850 ff.); and this developed into the Duden, which is the generally accepted standard today.¹

This Duden is a kind of linguistic chamber of horrors. In it we see the stuffed images of logic and classical grammar torturing a gaunt German language on the rack of grammatical analysis. Logic, for example, forces us to classify differently such sentences as *Das Wasser ist flüssig* and *Es regnet jetzt*, since the former is a 'logisches Urteil', whereas the latter is a 'Feststellung von Tatsachen' (342-3). More surprisingly, logic also explains the riddle of German word order: 'die gemeine oder gerade Wortfolge ... richtet sich im Deutschen nach streng logischen Gesetzen' (343). Classical grammar, in turn, twists German verb forms until they yield up the same six tenses as Latin and the same three voices as Greek. This does take a bit of torturing ('Um alle diese Formen zu bilden, muss sich unsere Sprache der Umschreibung durch Hilfszeitwörter ... bedienen'—89); but the editors of Duden have not been squeamish.

The most curious feature about this traditional analysis is the fact that, despite its authoritative standing, nobody is really satisfied with it: it fails too obviously in its appointed task. Several attempts at a new and better analysis have been made in recent years; the latest, and by far the most extensive and successful, is the book by Glinz here under review.

As a schoolteacher working with the 7th, 8th, and 9th grades, Glinz became acutely aware of the inadequacies of the traditional analysis. The first product of his dissatisfaction was his *Geschichte und Kritik der Lehre von den Satzgliedern in der deutschen Grammatik* (Zurich dissertation; Bern, 1947). This necessary but primarily negative study is now followed by a positive one, which was Glinz's Zurich 'Habilitationsschrift'. Its aim is, 'die Struktur unseres Deutsch so objektiv wie möglich zu erkennen und zu beschreiben' (11). Glinz believes (and we agree wholeheartedly) that a language must be analyzed in its own terms, and not on the basis of some external standard such as logic. What we must find is 'ein spracheigenes Kriterium' which will guide us in our observations, allow us to check our findings, and at the same time subject our findings to control by other observers. 'Ein solches Kriterium glauben wir nun in der Systemnatur der Sprache zu finden, wo uns nicht nur das Bezeichnende und das Bezeichnete [Saussure's *signifiant* and *signifié*], sondern auch das Zeichen selbst fassbar ist, nämlich in seinem *Funktionieren*' (52).

If we wish to observe the way in which the sign functions within the total system of a language, our first problem is: what body of materials shall we use? Ideally one should gather a large corpus and limit one's analysis to the materials in this corpus—trusting that it is large enough to contain everything worth analyzing. This is the procedure that Fries has used in his recent study of English structure.² Somewhat regretfully, Glinz rejects this method as impractical (53). As a substitute, he takes a relatively small corpus—mostly selections from

¹ Most recent edition of the Duden grammar: *Der grosse Duden: Grammatik der deutschen Sprache*, bearbeitet von Dr. Otto Basler unter Mitwirkung der Schriftleitung des Bibliographischen Instituts (Leipzig, 1935).

² Charles Carpenter Fries, *The structure of English: An introduction to the construction of English sentences* (New York, 1952).

Goethe's *Wilhelm Meister* (!)—and then proceeds to manipulate it by what he calls 'Systemerprobung'.

Glinz uses three methods of 'testing the system'. His principle method is the 'Verschiebeprobe', translatable as the 'test of rearrangement'. He takes a segment from the corpus and experiments to see in how many ways it can be rearranged. For example, the segment *Den andern Tag war eben alles wieder verschwunden* can be rearranged as *Den andern Tag war eben wieder alles verschwunden*, *Alles war eben den andern Tag wieder verschwunden*, etc. When carried out fully, the test reveals that this segment consists of six rearrangeable constituents: *den andern Tag, war, eben, alles, wieder, verschwunden*. Glinz calls these rearrangeable constituents 'Stellungsglieder' (87); they occupy a middle position between the word (which Glinz reluctantly defines as something written with a space on either side—81) and the clause (defined, by an admittedly rough criterion of intonation, as 'die kleinste Sprechereinheit ... der normal dahinfließenden Rede'—74).

This 'Verschiebeprobe' can be supplemented by the 'Ersatzprobe' or 'test of substitution' (similar to the American linguist's 'substitution frame'). In the segment quoted above we may substitute for *den andern Tag* such things as *nun, heute, am folgenden Tag, als Wilhelm wieder aufstand*; for *war* we may substitute such things as *ist, schien, blieb*; and so on. A still further test is that of 'Weglassbarkeit' or 'deletion' (the American linguist's 'principle of expansion', but in reverse). Three of the six rearrangeable constituents in the above segment (*den andern Tag, eben, wieder*) can be deleted, leaving *Alles war verschwunden* (or *Verschwunden war alles*) as a hard core of three non-deletable constituents.

These experimental tools, though objective, are somewhat rudimentary—particularly when applied to elements so inadequately defined as Glinz's 'word' and 'clause'. Glinz uses them with such skill, however, that he is able to reach an analysis very much like that produced by more refined and sophisticated methods.

A single example may suffice to show how skillfully Glinz works with his tools. By using the 'Ersatz- und Verschiebeprobe', Glinz arrives (96) at a clausal element which is 'nur einwortig und gleichartig ersetzbar', and which occurs (1) 'normal an zweiter Stelle' (*Er geht jetzt nach Hause; Jetzt geht er nach Hause; Warum geht er jetzt nach Hause?*), (2) 'nach gewissen Einleitungsstücken am Schluss' ([wenn, weil, etc.] *er jetzt nach Hause geht*), and (3) 'bei gewisser allgemeiner Inhaltsveränderung an der Spitze' (*Geht er jetzt nach Hause?; Geht er jetzt nach Hause, so ...*). Glinz calls this objectively analyzed element the 'Leitglied'; the customary term for it is of course 'finite verb form'.³

Glinz's further treatment of finite verb forms (99–113) comes close to being a model of analysis by form and function rather than by logic, meaning, and tradition. He first separates out the imperative. It differs from other forms (1) by the fact that it can occur indifferently in either first or second position (*Jetzt geh nach Hause = Geh jetzt nach Hause*); and (2) by the fact that it cannot

³ Like so many grammatical innovators, Glinz insists on inventing a whole new terminology of his own. This has certain advantages as he builds up his analysis during the course of the book. It also tries the patience of the reader.

occur in final position (the construction *wenn jetzt nach Hause geh* does not occur). The remaining finite verb forms are classified as follows (110):

		fest		anzu- nehmen	nur zu denken
		allgemein	vergangen	↓	↓
Einzahl	angesprochen	→ <i>bist</i>	<i>warst</i>	<i>seiest</i>	<i>wärest</i>
	sprechend	→ <i>bin</i>	→ <i>war</i>	<i>sei</i>	<i>wäre</i>
	besprochen	→ <i>ist</i>			
Mehrzahl	angesprochen	→ <i>seid</i>	<i>waret</i>	<i>seiet</i>	<i>wäret</i>
	sprechend oder	→ <i>sind</i>	<i>waren</i>	<i>seien</i>	<i>wären</i>
	besprochen				
Ein- oder Mehrzahl	höflich ange- sprochen	→ <i>sind</i>	<i>waren</i>	<i>seien</i>	<i>wären</i>

Glinz's terms 'allgemein' and 'vergangen' are strikingly similar to the 'non-past' and 'past' so dear to the hearts of many American linguists. Particularly noteworthy is the fact that he classes together 'sprechend' (1st person) and 'besprochen' (3rd person) everywhere except in the non-past singular, since it is only here that they differ formally. As far as your reviewer can see, the only mistake he has made is to set up a separate category for 'höflich angesprochen'. If we wish to classify formally (and Glinz certainly seems to wish this), then these forms of course belong in the category 'Mehrzahl, sprechend oder besprochen'—though we will now have to revise this label slightly.

This analysis of finite verb forms is a good sample of Glinz's work. Though he operates with premises and methods that seem at times hopelessly inadequate, what he produces is far and away the best structural analysis of German grammar published to date.⁴ Occasionally, of course, his inadequate methods are bound to lead him astray. In Chapter 5 they lead him to set up three non-finite verb forms: infinitive, present participle, and past participle. He recognizes that the three forms are not of equal functional rank when he writes (146): 'Im System der Vorgangsgefüge (umschriebenen Verbalformen) werden Rein- und Vollzugform [infinitive and past participle] auf eine Weise gebraucht, die der Artform [present participle] nicht zukommt.' If he had followed this line of reasoning a bit further, he would have seen that the infinitive and the past participle are examples of inflection (they play syntactic roles that cannot be played by any one-morpheme form); that the present participle is an example of derivation (it always plays a role that can also be played by a one-morpheme adjective, as Glinz demonstrates on p. 237); and that therefore the present participle cannot be ranked along with the infinitive and the past participle as a non-finite verb form.

⁴ It is of course true that Glinz's analysis is 'best' largely by default. As far as I know, the only other two attempts at anything approaching a structural analysis were both written for use with American students: Leonard Bloomfield, *First German book* (1928), esp. 276-315; and, relying heavily on Bloomfield, William G. Moulton, *Summary of German grammar, Dictionary of everyday usage: German-English, English-German* 2.467-504, ed. by J. Alan Pfeffer (1945).

Glinz's difficulty, of course, comes from the fact that he does not understand clearly the distinction between inflection and derivation. His reason for setting up present participles as a special verb form is the fact that they can be derived (with the suffix *-end*) from practically any verb, whereas other adjectives (in *-lich* or *-bar*, for example) can be derived from only a limited number of verbs. What he does not see is that this derivational fact merely establishes his present participle as a subclass of adjectives, not as a subclass of verbs. It would be equally senseless to set up ordinal numerals as a separate word class because they are all derived from cardinal numerals. On p. 297 Glinz very sensibly refuses to do just this; he calls them just plain adjectives—'reine Begleitartwörter'.

One thing that emerges with beautiful clarity from the pages of Glinz's book is the unevenness of form classes in German (and in all Indo-European languages?). We can make a division into (A) inflected words and (B) uninflected words or particles. Inflected words can be further subdivided into (A1) verbs and (A2) substantives; and substantives, in turn, can be still further subdivided into (A2a) nouns and (A2b) noun-modifiers. This class (A) and its subclasses (A1), (A2a), and (A2b) are all 'open classes': they are unlimited in size. The entire class (B) of particles, on the other hand, is strictly limited in size: it has only a few hundred members. Furthermore, though we can subdivide particles into various function classes (adverbs, prepositions, conjunctions), the amount of class cleavage is so enormous that we can no longer set up 'parts of speech' based on both function AND form. We must assume that a language is a system, and we must try to analyze it as a system; but it turns out to be a system that is very fuzzy at the edges. As Glinz very picturesquely phrases it, language is 'kein vollkommenes Gebilde ..., sondern eine vielfach überarbeitete, stets wieder geänderte, aber nie ganz umgebaute, immer nur geflickte und notdürftig angepasste, sich in jedem Teilhaber etwas anders spiegelnde Gemeinschaftsschöpfung der Menschen—eine Schöpfung, die sich zwar bei aller ihrer Unvollkommenheit doch immer nach gewissen Ideal-Richtlinien formt, stets auf dem Wege ist zu einem klareren, folgerichtigeren Bau, und die doch dieses Ziel, nach dem sie ewig strebt, ewig nie erreicht' (466-7).

Although Glinz's book is the nearest thing we have to a structural analysis of German, it suffers greatly from Glinz's almost complete lack of acquaintance with modern structural methods. He confesses (476) that he knows nothing about intonation, even though he has to use it in his definition of the clause (74). His first definition of the word ('die kleinste Einheit des Inhalts oder der Bedeutung'—79) would seem to refer not to the word at all, but to the morpheme. His later decision to define the word by using 'die übliche Abgrenzung, wie sie uns durch die Schrift gegeben ist' (81) is hardly acceptable in any serious linguistic study. Although he has sworn that he will base his analysis on form and function rather than on meaning, that old devil meaning creeps in again and again, particularly in the chapter on verb phrases.

To have reached, under handicaps like these, results that are by and large acceptable to the structural linguist, it is obvious that Glinz must have developed independently at least some of the analytical techniques used by structuralists. He uses the substitution frame as one of his basic techniques, though he never

really exploits it to the full. He develops at least partially the concept of immediate constituents. He uses the endocentric phrase, analyzed into head and attribute, though he never gets at the difference between this and the exocentric phrase. And he makes skillful use of the principle of minimal contrast to find out which syntactic features are meaningfully significant and which are not.

In terms of results achieved, structural linguists will be able to learn a great deal from Glinz. In terms of methodology, Glinz could have learned a great deal from structural linguists. The tragedy of his working in such scientific isolation is all the greater because of the impression one gets, again and again, that he would have grasped eagerly at the analytical techniques developed (in this country especially) during the last twenty years. In the very period when Glinz was working on his book (February 1942 to July 1948), there appeared in the pages of *LANGUAGE* alone at least a dozen articles which would certainly have interested him—from Harris's *Morpheme alternants in linguistic analysis* (1942) through Nida's *The analysis of grammatical constituents* (1948). Even aside from these more recent contributions, he would have gained enormously from an acquaintance with Bloomfield's *Language* (1933). The fact that Glinz did not know of any of these works is our loss as much as his.⁵

Glinz's book, then, is not nearly as good as it could have been. In one sense, we must condemn him for thus working under self-imposed handicaps. In another sense, however, we cannot help being filled with admiration as we watch him struggling against these handicaps and largely overcoming them. He has produced a work of genuine importance, and for that we must be grateful.

Ancient and mediaeval grammatical theory in Europe, with particular reference to modern linguistic doctrine. By R. H. ROBINS. Pp. 104. London: G. Bell & Sons Ltd., 1951.

Reviewed by HENRY M. HOENIGSWALD, *University of Pennsylvania*

The four chapters of this little book present grammatical theory among the Greeks, grammatical theory among the Romans, medieval grammatical theory, and a conclusion. One can well imagine that the series of lectures on which it was based (in the English Department of Birkbeck College) was very effective. Robins writes clearly, with a skillful use of repetition and an easy ability to keep the main lines of his argument in focus; and since he is guided by a thorough knowledge of current linguistics, his report on grammar in the Western world holds an interest to the linguist which such authors as Lersch (1841) and Steintal (1863) cannot offer. For data Robins naturally relies on these and other secondary sources; but the emphasis 'on those points that are relevant to linguistic and grammatical theory at the present time' (v) is his own, except that in the

⁵ Glinz seems also not to know George O. Curme, *A grammar of the German language* (2nd ed., New York, 1922; 7th printing 1952), despite the fact that it is the only full descriptive grammar of German ever written. This is all the more surprising since Curme's classic is regularly cited in the standard German handbooks. Cf. Otto Behaghel, *eschichte der deutschen Sprache*³ 3 (1928); Carl Karstien, *Historische deutsche Grammatik* 1.xxvi (1939); Adolf Bach, *Geschichte der deutschen Sprache*⁴ 212 (1949); Paul-Stolte, *Kurze deutsche Grammatik*² 43 (1951).

foreword he refers to an apparently unpublished Cambridge thesis by W. S. Allen on linguistic problems and their treatment in antiquity.

As a matter of fact the author is concerned less with what the ancient and medieval grammarians thought and how they came to think as they did, than with the extent to which their achievements anticipate present-day doctrine. This explains what would otherwise be remarkable omissions; the school of Pergamon, for instance, is nowhere mentioned, nor is the role of Remmius Palaemon at Rome (58) in supplanting Pergamenian with Alexandrian grammar. Medieval grammar is represented by Ælfric, Peter Helias, Alexander of Villedieu, the various writers *De Modis Significandi*, and little else; the nominalists, anti-peripatetic movements, and the like are hardly touched upon. Robins' position is one in which he must make value judgments, or at least must decide how far a given figure in history appears to have veered from the linear development of progress that leads up to our own day. If we accept the same position, we find that the judgments are well considered. There may be disagreement on matters of emphasis, or some reservations with regard to statements which have been made too sweeping in order to keep the main line of reasoning clear to the audience. When the author praises the Greeks for developing grammatical speculation without the benefit of extensive contact with other languages (6), one thinks of the possibility that the Stoic classification of verb tenses, with its unexpected subtlety (35), was the contribution of Zeno himself, whose mother tongue was Phoenician. On the whole, however, Robins is quite right in attributing the beginnings of grammar to 'the natural curiosity of the early Greek thinkers' (6). In fact, the Greeks are less guilty than he seems to think of lapsing from a strictly theoretical attitude toward language. It is true that our normative grammar, or rather the confusion between the task of describing a language and that of trying to make speakers change their speech, is unscientific and prescientific, and that its roots go back into the past (some of Robins' medieval material makes that very clear); but in antiquity it was hardly conspicuous before the period of the Atticistic movement, about the beginning of our era. Far more obvious than the seam between normative or utilitarian rhetoric and pure logic is the seam between pure logic and the kind of antiquarian activity which we associate with Alexandria. Though both can be found side by side at an early time, organic consistency was never quite achieved. In a way, the ancient dualism is reflected in our inveterate tendency to look upon phonology (the descendant of antiquarian concern with spellings, meter, and so on) and grammar in the narrower, quasi-logical sense of the word as two fields that profoundly differ in kind (93, with a different slant).

The author's approach is modified by an earnest endeavor to relate 'the study of grammar to its wider context in the intellectual climate of the times' (vi), but his dominant interest constantly takes him in the opposite direction. It is perhaps a little doubtful whether even the descriptive linguist (let alone the historian of ideas) fully shares this interest. Rightly or wrongly he may well take the position of the Arab conqueror: if the books of the infidels agree with the Koran, they are not needed; if they disagree, they are evil; so let them be burned. What is more, if the history of grammar is presented to the present-day worker in the field as a

collection of errors and truths from his point of view, he will be suspicious of truths found mixed with so much patent error. Natural scientists today are not all interested in the history of their disciplines; but those who are show just as much concern for conceptions that were later abandoned as for achievements that have remained. Nothing could be more revealing than to compare Robins' comment on a certain passage in Quintilian (1st century A.D.) with that of Jacob Wackernagel (*Vorlesungen über Syntax* 19 f.) on the same passage. Quintilian argues for a seventh case in Latin, and for a fifth in Greek, on the ground that Latin ablative and Greek dative forms often have a special instrumental meaning. To Wackernagel this is a remarkable achievement in that it anticipates one of the results of modern comparative grammar, to the effect that the ancestor of both languages once had a separate, formally characterized instrumental case, which fell together with the dative-locative in Greek, and with the ablative-locative in Latin. To Robins it is a fall from grace, in that Quintilian introduces without warrant a semantic criterion into an otherwise fairly consistent formal classification (59 f.). No doubt Wackernagel's praise is undeserved, while Robins' criticism would certainly be appropriate if Quintilian were writing now. Quite possibly it is just even if one takes into account the standards of Quintilian's own time; but the general orientation of Robins' work somewhat obscures this. At another point (82 f.) Siger of Courtrai (ca. 1300 A.D.) is said to have anticipated de Saussure's theory of the linguistic sign, with the added remark that 'the purely philosophical implications of this [viz. Siger's formulation] need not trouble us.' Granting that the resemblance is striking, it suggests that there is more point in looking for unconscious dependence of modern scholars on early speculation than in identifying instances of anticipation. That the celebrated pair *signifiant-signifié* reads like a translation of the Stoics' *σημαίνον-σημαινόμενον* (mentioned on p. 26, with the accent to be corrected) is perhaps an asset for the Greeks, but more probably merely a liability for modern linguistics. Without imagery of this kind, so dear to the more popular forms of philosophizing in the West through the ages, it might have been easier to come to a linguistically satisfactory concept of meaning. According to the Ionian physicians, speech sounds are influenced by climate, a theory that made good sense beside the other assumptions in their philosophy of nature. Men like Collitz and Prokosch might have been somewhat more cautious in mistaking the familiarity of this motif for self-evidence if they had been aware of its explicit history.

It would be folly to blame Robins for doing what he has chosen to do (and has done excellently) instead of trying to write the real history of Western grammar from the Greek beginnings to the Renaissance. There are many reasons why this cannot be done at present: one, of course, lies in the recent history of linguistics itself, which has just begun to make possible an infinitely more detached attitude toward its earlier past than was conceivable when Lersch and Steinthal wrote; other factors are the sheer amount of historical research (e.g. on Varro, on the Middle Ages) that waits to be interpreted, and finally the stimulation emanating from the history of science and the history of ideas as now understood. Pending that, we may be grateful to have Robins' informative little work.

Beiträge zur Etruskologie: I. Silbenpunktierung und Silbenbildung im Altetruskischen. By FRIEDRICH SLOTTY. (Bibliothek der allgemeinen Sprachwissenschaft; 3. Reihe: Darstellungen und Untersuchungen aus einzelnen Sprachen.) Pp. xvi, 207. Heidelberg: Carl Winter, Universitätsverlag, 1952.

Reviewed by HENRY M. HOENIGSWALD, *University of Pennsylvania*

In certain early Etruscan inscriptions, the letters are often accompanied by dot-like marks. The practice is also known elsewhere in ancient Italy, notably in the Venetic area, where it must have spread from the Etruscans along with the alphabet itself. The dots are clearly not word separators (those inscriptions normally lack the word-dividing punctuation which became so common in Latin): they mark word-initial vowels and syllable-final consonant groups or consonants. This at least was Deecke's and Thurneysen's interpretation of the Venetic inscriptions sixty years ago, when the remaining evidence was virtually unknown. Conway refused to accept it; but now, after Vetter's thorough re-examination of the whole matter in the light of the increased material, no serious doubt is possible any longer. One must only ask why such a system should have developed in the history of the script, and what, if anything, it teaches us with regard to the sounds of the languages represented.

In spite of Slotty's argument (62 ff.), it still looks as though the syllabic puncts were relics of pre-alphabetic writing. Kretschmer, in an article unknown to Slotty,¹ thinks of an old system evolved in Crete and the Near East long before the so-called invention of the vowel signs by the Greeks; he points to the *virāma* of the Indic alphabets not as a parallel but as a descendant, and to many features in the history of the Semitic scripts. Perhaps not many Semitic scholars will follow him in this. The more conservative view connects the syllabic dots in Italy with a hypothetical syllabary, perhaps of Minoan extraction, used (one would have to think) by the Etruscans before they received the Greek alphabet. The dots were kept for some time even after they were no longer needed, either from sheer inertia, or more probably because they had a function in the teaching and learning of literacy. For the Veneti at least, according to Beeler,² this is plausible; and it was there that syllabic punctuation continued in existence, while it was soon given up in Etruria and Campania. It is true that the very oldest Etruscan inscriptions do not show it, but this does not prove it an innovation rather than a relic. Of the eleven inscriptions listed as oldest by Ducati (*Etruria antica* 66 f.), only four, three of them extremely short, come from a site where so far as we can tell it was at any time in regular use (Caere). If we know anything about the alphabet in early Italy it is that local differences at the beginning were far greater than later, and that every town, certainly every region, had its own epichoric script. From this, many scholars have argued that the Etruscans acquired the art of writing in small groups before they settled in their historic seats. By and large, then, in those localities where syllabic punctuation exists at all, it is the oldest inscriptions that tend to show it. Absolute neatness

¹ *Minos* 1.7 ff. (1951).

² *The Venetic language* 11 (1949).

cannot be expected, because—unlike the Veneti—the Etruscans apparently did not consider the puncts an integral part of spelling: even at one place in one period, marked and unmarked texts are mingled. What is more serious is that there is no direct evidence for a syllabary in Italy, nor for the use of punctuation in the only syllabary known in the Mediterranean, the Cyprian. Still it is worth pointing out, as this reviewer has done elsewhere,³ the rather specific resemblances between the kind of syllabic script one would reconstruct from Etruscan syllabic punctuation and the kind of script that the Cyprian syllabary actually was. Syllable-final consonants and consonant groups are dotted in Etruscan; in Cyprian they can only be written with signs that stand primarily for consonant plus vowel (*e*). In other words, Cyprian could use a virāma or minus-vowel mark of some kind. Syllable-initial consonant groups (in the clearest case word-initial) are not marked in Etruscan; in Cyprian there are special signs for some though by no means all such groups. J. F. Daniel thought that at least one of them (*sdo*) might be pre-classical, i.e. go back toward if not into Minoan times (but see below). Word-initial vowels are, on the whole, marked in Etruscan; the Cyprian signs for vowels without consonants seem to have been secondary inventions. This would point to the use of (say) dotted *ve* for *u* on the precise analogy of dotted *te* for *t*. The sequence *ia* is in archaic Etruscan often written *īia*; in neither case is *a* dotted (as one might expect it to be in *ia* with true hiatus, parallel to word-initial *a*); in Cyprus the special signs *ya*, *ye* are used after the *-i* signs (*i*, *pi*, *si*, etc.), a practice which may be responsible for the *ua* spellings on the alphabetic (not syllabary!) records of nearby Pamphylia. Of course this is not to suggest that the hypothetical Etruscan syllabary was derived from Cyprus. Both may reflect some common post-Minoan model.

On the other hand, it may all be a mirage. It is a curious fact that initial *i* is dotted only in southern Etruria (i.e. in Vei, Caere, and a few scattered instances elsewhere in the extreme South) and in the Venetic script; in Campania it is unmarked (86 f.). Slotty suggests that the omission served the purpose of word division in a roundabout way: since the diphthongs like *ai* at the end or within a syllable have their *i* (a consonant for the effects of the rule) dotted, an *i* not so characterized would indicate syllabic value, hence the beginning of a new word after a final vowel. The need would arise only for *i-*, not for *a* and *e*, and not for *u*, since the diphthong is always spelled *av* (with dotted digamma), never *au*. This use would have spread to position after word-final, therefore syllable-final, and so to dotted consonants where no such marking was needed. As Slotty presents it, this seems a very narrow base to support his argument, but the argument is much stronger than he seems to know: it holds not only for word-initial syllabic *i* (i.e. historically a *y*-sign without its inherent vowel?) but also for the word-initial consonant clusters (*st-* etc.) the first members of which remain mysteriously without puncts. There is only the small discrepancy of locality. We must assume that the invention at first affected the ordinary consonants, while *i* was added in Campania after colonization. Whether the near-lack of initial non-syllabic *i* in Etruscan (which Slotty hardly succeeds in denying) has something to do with all this, too, need not be discussed here (166 ff.). If Slotty

³ AJA 56.174 (1952).

is right, and if we are right in extending his train of thought, the existence in the Cyprian syllabary of a few signs for *sd* plus vowel and *ks* plus vowel may be only a coincidence. It is somewhat suspect anyway, because these two clusters are precisely the ones for which there happen to be special characters (zeta, xei) in the Greek alphabets. Perhaps we have here no more than a late effect of Greek alphabetic writing on the Cyprian syllabary.

Cyprus or no, there is still every reason to regard syllabic punctuation as a fossil. In Beeler's words,⁴ it is 'of some service ... in the division of words ... But even here the punctuation is not generally useful, for it is of no value, for instance, when one word ends with a consonant and the next begins with one, and when the preceding words ends with a vowel and the following begins with a consonant ... In respects other than that of word division the Venetic puncts indicate nothing which is not already obvious ... It is evident, then, that such a system was invented neither for Venetic nor for Etruscan.' Slotty takes a completely different view. As reported above, he believes that syllabic punctuation was invented in (southern) Etruria around 550 and kept in use for about a hundred years for the express purpose of recording the boundaries of chest-pulse syllables in spoken Etruscan. In his opinion the Etruscans were led to this unique achievement by a change in the nature of the stress accent, which became strongly word-initial in the 6th century; 'such an accentuation has a secondary accent on the third syllable as a natural result. Such an accentuation cannot but stimulate and enhance the feeling for rhythm to an extraordinary degree. And in my opinion it is this increased feeling for the rhythmic nature of the sentence that results in the marking of the syllables as rhythmic entities' (65). A century later, syncope of vowels catches up with the strong initial stress, the rhythm of the sentence collapses, and syllable boundaries are no longer marked. We, too, have an experience of the crests and troughs of breath pressure in our respective native languages, but it is not sufficiently strong to make us express it in writing. The Etruscans are credited with a fine ear and with strong rhythmic feeling, the latter evident also in the representations of dancers on the walls of the Tarquinia tombs (64).

This is no more credible than was Conway's notion of the Venetic punctuation as a system of word accentuation. As a rule, speakers use the available alphabetic resources to render the distinctive sound features of their language as best they can. In the Western tradition entire areas of distinctive sound feature, such as stress and pitch, have remained all but neglected, owing mainly to the historical accident which gave us procedures to write consonant and even vowel segments but little more. Nondistinctive features may occasionally be committed to writing where symbols are available, though even for that really good examples are doubtful. But only phoneticians invent devices to write down matters as automatic (even for Slotty, or he would not claim to know just where to expect the dots) and as elusive (perhaps also as ill-defined both phonetically and phonemically) as chest-pulse troughs. Can a single graphic development even remotely approaching such a feat be found anywhere in history? And if the Etruscans had really chosen to adorn their pottery with phonetic recordings of

⁴ Op.cit. 12 f.

this sort, would they not have invented a simpler system than one which required them to mark, for example, all the members of a syllable-final consonant group with dots? The chances are that syllabic punctuation no more renders anything phonologically significant that might be called a syllable than does the akṣara of the Indic alphabets: there, a sequence such as *arta* is, graphically speaking, *a rta*, whereas the syllable boundary—as suggested by such factors as accent distribution, prosody (in meter), and no doubt also the chest-pulse boundary if it can be located—lies after the *r*. On a Campano-Etruscan bowl the name *maifnas* (58) has dots at the *i*, the *f*,⁵ the *l*, and the *s*; this tells us, possibly, how the scribes handled a difficult sequence in which the last syllable (in the canonic sense) began undisputably with *n*, but where there was no precedent for the remainder of the word (accumulations of consonants of this type being the result of vowel syncope and still infrequent in punctated inscriptions). It certainly does not tell us anything about the chest pulses in its pronunciation.

Aside from this theory, Sloty's book is a marvelous contribution to Italic epigraphy. His material for Etruria and Campania is as complete as anyone could make it, and is based on new readings both by himself and by Vetter. Other sites are excluded, and wisely so, from the treatment. The inscriptions are transliterated with full critical apparatus, including Vetter's valuable new reading of the Capua tablet, one of the longest Etruscan texts. Various types of syllabification, linguistic or graphic, are discussed, with the interesting result that the Capua tablet tends to treat certain consonant groups as straddling the boundary where other texts, in Sloty's opinion less conservative because less formal, treat them as syllable-initial—a trend which, if it really reflects a difference in pronunciation, would parallel the correptio Attica in Greek prosody. By all odds the most valuable part of the work is the second, consisting of twelve excursus, among which there is a thorough presentation of new facts on word division implied in the syllable marks, and a word list for the Capua tablet. There could be endless discussion of individual readings and other epigraphic detail, mostly in a spirit of agreement and admiration. This is an excellent and most important book.

Morphologische Neuerungen im altindischen Verbalsystem. By M. LEUMANN. (Mededelingen der Koninklijke nederlandse Akademie van Wetenschappen, Afd. Letterkunde, NR Vol. 15, No. 3.) Pp. 51. Amsterdam: N.V. Noord-Hollandsche Uitgevers Maatschappij, 1952.

Reviewed by LOUIS RENOU, *Sorbonne and Yale University*

The time is past when one could hope to use Indo-European to explain many developments peculiar to the historical period; today innovations are considered more important than survivals. Such is the case with the *r*-endings in Indo-Iranian, especially in Vedic, which have evolved in a manner altogether different from other Indo-European languages. Leumann is right in trying to determine, for every particular case, the morphological environment—to reason in terms of the general system to which all these endings are tributaries. The field is happily

⁵ The dotted *f* need by no means be a 'mistake', as Sloty thinks.

chosen to illustrate the unusual fluidity of the language in the Vedic stage: many formations were roughed out, tentatively expressed, or abolished, whose starting-point we cannot reconstruct without a careful examination of all parallel structures. Thus Leumann is right in saying that the ending *-ran* developed from *aśeran* and *aduhran*; these forms again are connected with *śere* and *duhre*, and farther with the 3d sg. *śaye* and *duhe*, which came to be involved in the domain of the perfect for reasons satisfactorily explained by the author. This *-ran* is nothing more than a normalization of *-ra*; but the alternative form *-ram* is a phonetic deviant of *-ran*, starting from *aśgram*. In the sigmatic aorist, the extension of *-ur* is correctly interpreted in connection with the paradigm *āsa/āsar*. (Incidentally, the 1st sg. *āsa* beside *āsam* is assigned—perhaps too boldly—to the imperfect, and is joined to Hom. *ῥα*.) Leumann is equally right in rejecting the conception of perfects in *-ur* without reduplication (except of course for *vidur*); thus, he convincingly analyzes *mandur* as resting on a weak stem *mam-d-* corresponding to the strong stem *mamad-*.

The latter and shorter part of the booklet deals with certain other problems of Vedic morphology: the origin of the 'precative'; the origin of the *-siṣ-*aorist; the prototypes of the desiderative (where Leumann sees, in the relation of the present stem *dad-* to the desiderative *dītsati*, the model for the extension of the type *pītsati*, *mītsati*, etc.); and finally the dependence of the *-sa-*aorist on the *-s-*aorist, concerning which Leumann clarifies Wackernagel's views by noting the importance of the form *adhukṣata*, which was fundamentally ambiguous.

The entire demonstration is managed with profound knowledge and penetrating insight.

Remarques sur la place du verbe dans la phrase active et moyenne en langue sanscrite. By J. GONDA. Pp. 86. Utrecht: A. Oosthoek, 1952.

Reviewed by LOUIS RENOU, *Sorbonne and Yale University*

Studies of word order have been less numerous in the domain of Sanskrit than in that of other Indo-European languages. Probably it is the prevalence of poetical texts that has hampered research, since workers may have feared that such investigations would have no precise result. The Dutch scholar Gonda, who has already rendered so many services to Sanskrit grammar, did not hesitate to face the difficulty. The book before us, based on extensive selections from the *Ṛgveda* to the *Rājatarāṅgiṇī*, not only does not avoid the use of metrical texts, but contends that they have much to teach us, even despite their apparent freedom and irregularity of order.

Gonda's study of course confirms our knowledge that the personal verb in Sanskrit is most commonly at the end of the sentence; but its tendency to occupy this position is opposed by others, so that it cannot be defined as normal without qualification. Position at the beginning of a phrase depends on the nature of the verb (for instance, if it is an imperative), or on the connection with a preceding phrase (for instance, if the phrase concerned is an apodosis), or on the nature of the phrase itself (for instance, if it is interrogative or 'affective'). The most delicate problem is offered by the many instances where the verb is medial in

its phrase, and where we must first decide whether it is treated as a kind of enclitic. Gonda shows that medial position of the verb is attested principally where there is inversion of the subject (for instance, the epexegetis of a proper noun, of the type *tad wāca Gārgyaḥ*), or where the phrase includes an enlargement. He draws attention to subordinate clauses, in which (at least in prose) the final position of the verb is almost constant.

The treatment of the R̥gveda is perhaps too short. Probably the question needs to be examined in correlation with the place of the separate preverb, since the verb stands to this as a kind of satellite.

In a very elaborate last chapter, Gonda takes up the whole problem anew from the Indo-European point of view. He concludes, with good reason, that most of the tendencies attested in Sanskrit are present also in the cognate languages, in some of them in fact until their modern stages. It is true that the end position of the verb has been largely given up; but one of the aims of the present work is precisely to show how this development was foreshadowed in earlier times. Since Gonda has succeeded in escaping the subjective prejudices so common in dealing with these elusive problems, his work generally carries conviction.

La dialectología griega como fuente para el estudio de las migraciones indoeuropeas en Grecia. By F. RODRIGUEZ ADRADOS. (Acta salmanticensia: Filosofía y letras, Vol. 5, No. 3.) Pp. 73. Universidad de Salamanca, 1952.

Reviewed by CARL D. BUCK, *University of Chicago*

The author first reviews the opinions of linguists and historians on the subject indicated by his title. (Among these, the present reviewer rates as a conservative, if not a pessimist.) The bulk of the work consists of a detailed discussion of the various isoglosses. They are the same that have been discussed by others. But the author believes that he is employing an improved, more rigorous method of evaluating them, by a classification into archaisms, innovations, and selections, and by the exclusion of all non-linguistic evidence. As to the last, few scholars are willing to ignore historical tradition when it fits in with the linguistic evidence so well as it does in the case of Thessalian and Boeotian. As to the three classes, the fact is that in each of them there are some isoglosses that are significant and some that are not. Even agreement in archaisms, though in general they are less impressive than the innovations, does not lack significance when it lines up with other features, e.g. the persistence of the old *δίδωσι* etc. vs. the innovating *δίδωσι* etc. In fact, in Thessalian and Boeotian, forms like *δίδωσι* are—from the viewpoint of these dialects—innovations representing a West Greek overlay on an Aeolic foundation. Incidentally, the author has missed the most important discovery of recent decades, the appearance in the inscriptions unearthed at Cyrene of forms like *παῖσα*, *ἐκοῖσα* (nom. sg. fem. pple.) *ἐμμένουσι* (dat. pl. pple.), agreeing with Lesbian *παῖσα* etc. in the treatment of secondary medial *νσ*, and so proving an Aeolic element in this Doric dialect—a striking match for the Lac. *Ποιοῖδαν* reflecting the pre-Doric form shown in Arc. *Ποιοῖδαν*. The new Cyrenean inscriptions also show that the old view of *τελεσφορέντες*, repeated in the present book (61), as an example of the unthematic inflection of contract verbs, is wrong.

It is rather a phonetic matter, ϵ from eo before vr , as shown by $\Lambda\epsilon\upsilon\tau\iota\chi\omicron\varsigma$ = $\Lambda\epsilon\omicron\upsilon\tau\iota\chi\omicron\varsigma$, and the 3d plural Doric future $\acute{\epsilon}\kappa\tau\iota\mu\alpha\sigma\acute{\epsilon}\nu\tau\iota$.

The author's conclusions are summarized on pp. 67 ff. The last wave of migration was that represented by the West Greek dialects. On this there is no dispute. Of the early waves, the priority of the Ionic is assumed without reservation. Actually this is a not unlikely hypothesis, now widely favored; but the arguments for it are more historical than linguistic. There followed the waves which the author calls Aeolic, using this term in a wide sense to include what became Arcado-Cyprian ('el primero de los dialectos eolicos, el arcadio-cypriota'). Here there is a real question of classification involved. Arcado-Cyprian has several striking characteristics in common with Aeolic, which indicate a prehistoric contiguity in the north, very probably antedating the invasion of Greece. There was no doubt some common development, but by no means complete unity; for Arcado-Cyprian lacks several of the most characteristic features of Aeolic, and moreover exhibits some features common to Attic-Ionic, as well as special traits of its own. If we classify the dialects as we actually find them, it is the part of realism to recognize among the older dialects three coordinate groups: Attic-Ionic, Aeolic, and Arcado-Cyprian. This is my view and that of many others.

With all my reservations, I welcome this work as further evidence—added to Tovar's article in *Emerita* for 1944—of a critical interest in the Greek dialects among Spanish scholars.

Le cheval dans l'Iliade: Suivi d'un lexique du cheval chez Homère et d'un essai sur le cheval pré-homérique. By EDOUARD DELEBECQUE. (*Études et commentaires*, Vol. 9.) Pp. 251, offset. Paris: Librairie C. Klincksieck, 1951.

Reviewed by JOSHUA WHATMOUGH, *Harvard University*

It appears that the word $\iota\pi\pi\omicron\varsigma$ 'horse' occurs 417 times in the Iliad: 25 times in the singular, 16 in the dual, and all the rest in the plural. In the sense of 'chariot' it appears only in the dual and plural. Moreover, Homer's heroes do not ride horses in the Iliad, unless in 10.513 (according to a dubious interpretation of that line in a book of admitted peculiarity and demonstrable late composition), even if equitation, clearly mentioned only once in the Iliad (15.679) and that in a simile, was known to someone who had a hand in its composition at a comparatively late date; even the ancient critics were aware of the difficulties both at 15.679 (Aristarchus) and in Book 10 as a whole (schol. T). The use of the simile is far from proof that the 'poet' had 'seen horse-riding with his own eyes' (Delebecque 76, cf. 112); a modern poet might invoke a comparison with life on the moon or on Mars, but he has never been there.

The matter is trifling, and of no interest to linguists except those few who happen to be interested in the Homeric question. The only other topic touched upon by Delebecque that is likely to interest readers of *LANGUAGE* is what he calls the 'pre-Homeric' horse. It is clearly established that 'the horse was the chief means of locomotion by which the expansion of the early Indo-European languages was accomplished, and for this reason some Indo-Europeanists show as much interest as zoologists in the evolution and domestication of the horse.

Delebecque has a bibliography that runs to seventy-nine items, to which I can add from my own collections nine others: A. D. Fraser, *The Greek cart-horse*, *Classical journal* 39.447 (1936); A. Heiermeier, *Westeuropäische Heimat und Namen des Pferdes*, *Padeia* 6.357-75 (1951); M. Hilzheimer, *Evolution of the domestic horse*, *Antiquity* 9.133-9 (1935), followed by R. H. Lane, *Waggons and their ancestors*; A. Hörschemeyer, *Die Pferdezucht im klassischen Altertum* (Diss. Giessen, 1929); E. Raucq, *Contributions à la linguistique des noms d'animaux en indo-européen* (Antwerp, 1939); A. Sakkas, *Le cheval dans la Grèce antique* (Paris, 1934); M. Wolff and D. Opitz, *Jagd zu Pferde in der altorientalischen und klassischen Kunst*, *Archiv für Orientforschung* 10.317-59 (1935-36); and above all Ridgeway's *Origin and influence of the thoroughbred horse* (Cambridge, 1905).

Greece proper is a poor country for horses; Thucydides noted the difficulty of feeding a large number of horses in Attica (7.27). There is only one word for horse in Greek, and it has features (the rough-breathing and the initial iota) which point to a non-Greek source. Contrast with this the variety of names recorded in western Europe: not only *equus* (and *Epona*, *ebul-*) but *caballus*, *marcus* and *marca*, *mannus*, *ginnus*, *ueredus* (-arius) and *paraueredus*, *celdo*, *calliomarcus*, all of which I cull from *DAG*. But by the date from which even the earliest recorded of these comes, the horse had been put to many other uses than drawing a chariot. As for *marc-*, limited to Germanic and Keltic among Indo-European languages, the Mongolian *mori* 'horse' is doubtless the same word. But which is borrowed from which?

Gli antichi Italici, 2d ed. By GIACOMO DEVOTO. (Collana storica, Vol. 29.) Pp. 356. Firenze: Valecchi Editore, 1951.

Reviewed by JOSHUA WHATMOUGH, *Harvard University*

Since this work, first published in 1929, has reached a second edition, it evidently has readers. Not that there is much new to be said, except by the archaeologists, whose minds are changed too often to inspire much confidence in their opinions. Devoto, whose performance echoes his name, does indeed mention one new Umbrian inscription (38), but he seems not to know about another (see *HSCP* 50.89-93 [1939]). On the linguistic side he has kept abreast of new work in all its ramifications. For example, he cites with approval (134 n. 21) Bonfante's view that the Latin system of personal names is essentially Indo-European (*Mélanges Marouzeau* [1948]), a view advanced simultaneously by Pulgram (*HSCP* 58/9.163-87 [1948]); and (98 and n. 3) the view of Merlo, Holmes, and Pulgram (*Lg.* 25.247 [1949]) that sees Etruscan influence in 'l'aspirazione fiorentina'. There is a good survey of work in the Italic field by Büchner and Hofmann, *Lateinische Literatur und Sprache in der Forschung seit 1937* (Bern, 1951), who deal, like Devoto (very briefly, in his Chapter 7) with the problem of a native literature, a matter considered also by Bardon in his recent book *La littérature latine inconnue* (Paris, 1952), reviewed by me in a forthcoming number of the international journal *Erasmus* (Basel).

My *Foundations of Roman Italy* (London, 1937) also attempted to combine

the evidence of archeology, anthropology (chiefly physical), historical tradition, folklore, and language—subject always to adherence to a strict rule that they refer to the same people, in the same place, and at the same time. Reviews in archeological journals protested that too much attention was paid to linguistics; in linguistic journals that too much attention was paid to archeology—from which it may be concluded that the proportions were about right. But if I were writing that book now—and this is what I miss most in Devoto's pages—I should attempt to apply the technique of social anthropology, especially the search for a cultural pattern as a correlate of the linguistic, along the lines of the discussions at the Linguistic Institute last summer in Bloomington, which proved so fruitful. Nothing of this appears in Devoto's *Gli antichi Italici*.

Devoto moreover interprets *Italic* in the strictest sense. But even the *Italici* of the archeologists are not precisely the *Italici* of language. Both disciplines exclude Latin, and everything that Conway and I called 'Prae-Italic'. Further, Devoto holds the hypothesis, by no means universally accepted, that the two subdivisions of what is commonly called Italic, namely Latin-Faliscan and Oscan-Umbrian, are somewhat distantly related. As for Italo-Keltic, a unity still maintained by Ernout in the most recent edition (1951) of his and Meillet's Latin etymological dictionary, it naturally finds no place. Instead we have the dubious archeological concept of 'Proto-Italici'; the 'Proto-Kelts' of Childe and Pokorny are not, I think, even mentioned. It will be clear what I consider the chief weaknesses of Devoto's book.

On the linguistic side, subject to the reservations already made, it is excellent. There is now general agreement about the interpretation of Italic texts and Devoto is an accurate and readable reporter, with enough originality to stimulate, but not so much as to shake the confidence of his readers. He has a sure command of his materials and he has devised an interesting plan for his book, which begins with an entertaining account of discovery, starting in 1456 when one Paolo di Gregorio Scano was bargaining with representatives of Gubbio for the purchase of the Iguvine Tables, at least the seven surviving out of the nine that had come to light in 1444. Then the position of the Italic dialects is discussed in relation to Indo-European and to their neighbors in ancient Italy. Once the 'Proto-Italici' are disposed of, we take up in succession the historical *Italici*, their expansion, the Italic documents and their alphabets, art, culture, religion, political institutions, and the part they played vis-à-vis Rome, ending with the Social War (91-88 B.C.). Had the 'Socii' been victors in that fierce contest, the Romance world would now be speaking languages descended not from Latin, but from Oscan (perhaps rather North Oscan). Even six years later, when the Samnites were thrust back at the battle of the Colline gate, hatred of Rome still raged. The weapons and armor of the Samnites engaged in that bloody battle show a high quality of artistry that reveals the central Italic tribes no inferiors of the Romans in material civilization. A thoroughgoing excavation of the Italic capital of Corfinium, and of its vicinity, whenever it is made, is likely to be rewarding—even dialect inscriptions may lie not far below the surface.

A few details. 13: *Grecchio* is the accepted spelling. 42: Greek *-ῆ-* is not the representative of *η* but of *ῆα*, or, according to recent theory, of the reduced

grade of *en-eH*₁ (e.g. *κασιγνήτος*, cf. *γενέτωρ*). 47: The vocable *pur*- 'fire' is not unknown to Latin, for *purgare* is 'purify with fire' in the first place. 50: If Beeler is right about the proper classification of Venetic, and I think that he is, then the testimony of that dialect greatly strengthens the interpretation of the inscriptions of the Val Camonica as Latinian (Proto-Latin?), not Italic. 69: Footnotes 18 and 19 are badly pied. 72: Another view of Lat. *sepeliō* would appeal to Lepontic *pala*, if not to Umbrian *pelsa*-, for Vedic Sanskrit *saparyāti* means not 'inter', but 'honor'. 118: For Umbrian *casilos* compare the Lepontic personal name *kasilos*, PID 327. 129: The ethnicon *Aequi*, in view of the Greek spellings quoted by Devoto, and of the local names *Aecae* (Daunii), if not *Aecus Tuticus*, may be presumed to have *-ky-*, not *q^u*. 151: Read Athenaeum (not *-aeun*). 177. The Oscan *ettuns*-inscriptions find a remarkable parallel in certain Latin inscriptions of the Rhineland, which I expect to discuss in a later number of this journal. 222: Devoto, if I understand him aright, appears to count the Oscan divine epithet *fragiūt* (Cumae) a variant of *flagiūt* (Capua); I agree; Hofmann (LEW³ 510 s.v. *flagrō*) would set up **bhereg-* beside **bheleg-* from **bher-* and **bhel-* (of identical meaning). That the initial cluster *fl-* was unstable in Oscan is clear from *fluusasiats* beside *fluusat* at Agnone; Umbrian had no forms with *fl-* at all.

An Oscan text that has given much trouble is the *memnim*-curse from Capua, ID 131. The rule is that *o*-stems have nom. sg. *-s*, acc. sg. *-ūm*; but *io*-stems have nom. sg. *-is* in the prae nomina, *-iis* in the gentilicia, yet acc. sg. *-im* (prae nomina) and *-iium* (gentilicia). Whatever the reason for this variation may be (the best explanation is that given in ID 2.470-2), the distinction between prae nomina and gentile names is clear. Hence the only way to interpret ID 131 is to adhere to the rule and to bring together the pairs of names (e.g. *plasis ūppiis*, *bivellis helleviis*, and so on throughout the text) which were separated by the composer of the curse, presumably for magical purposes. Then, in line 7 *hukis* and *uhtavis* are two separate persons (Lucius and Octavius) and *pūtad* is construed twice over, i.e. with each in turn. In this way good sense is obtained without violation of grammar.

Le lingue dell'Italia antica oltre il latino, con due carte geolinguistiche ed una tavola degli alfabeti. By VITTORE PISANI. (Manuale storico della lingua latina, Vol. 4.) Pp. xvii, 354, with folding chart and two maps in pocket. Torino: Rosenberg & Sellier, 1953.

Reviewed by JOSHUA WHATMOUGH, *Harvard University*

The copy which I am reviewing was posted on 16 December 1952; the preface is dated 4 November 1951; the publisher's date 1953 is therefore bogus. The two maps, usefully placed under a flap, are marked, truthfully enough, *Carta linguistica dell'Italia antica I* and ... *II*; they give the dialects, tribes, some regional names (including the bastard *Bruttium*), and provenances of inscriptions; the high-falutin 'geolinguistica' label of the title page (what else can a language map be?) gives out an unpleasant echo of 'geopolitics' that would have been better avoided. The two tables of alphabets are pirated, with hardly any change and with

no acknowledgment whatever, from *PID* and *ID*. The book shows little more originality elsewhere, except when it shows too much, e.g. by inventing a second edition of Buck's *Grammar of Oscan and Umbrian*, fathered upon a publisher in Chicago, 1933. What the author has done is to skim the cream off von Planta, Conway, Buck, Whatmough, Devoto. Thus he gets both Italic and Prae-Italic, as well as some Etruscan, under one cover. There is a brief sketch of the grammar of each dialect, a translation into Latin, full commentary to each text, and alphabetical indexes of the words discussed.

The best compilation for use with students was Hermann Jacobsohn's *Altitalische Inschriften*, now out of print (Lietzmann's *Kleine Texte*, Vol. 57; Bonn, 1910; anastatic reprint, Berlin, 1927). Too much help, and in particular translations, should not be given to students. Pisani gives both. The merit of his compilations is that, like Jacobsohn's, it covers the entire area of ancient Italy.

P. xiv: I know a G. S. Conway, the son of R. S., but no C. S.; add *Dialectorum italicarum exempla selecta*, also by R. S. Conway (Cambridge, England, 1899). The worst of dating 1953 a book that you wrote two years earlier is that you are left with the second edition of Ernout and Meillet instead of the two-volume third edition of 1951. xv: Lindsay's 1930 edition of Festus should have been identified as part of Volume 4 of the British Academy's *Glossaria latina*. 2: The reader who does not speak Italian is likely to be misled by the information given here on the pronunciation of OU *z* in the native alphabets. 3: When *h* is used as a mark of hiatus it is between unlike vowels, not 'fra vocali'. 4: The bald assertion that *p* in *eitipes* is an unsuccessful attempt to write *f* rests upon an unsuccessful attempt to explain *eitipes*. 6: If *Pakim* stands for **Pakiom*, how are we to account for *Kluwatium*? Samprasāraṇa does not take place before -*m*, or even syncope (O *hūrtūm*, U *saluom*). 7: The first syllable of *dīviat* has *ī*, not *ī*; and the correct translation is *diuinae*, not *Diae*. As for *pīstiat*, the spelling seems to have been influenced by that of *pīthiūt*. 8: *uesclis* (Volscian) and *uesclir* (Umbrian) are poor evidence for Umbrian *y*- from *l*-. This critique could be easily continued page by page. Let these examples suffice. But I cannot refrain from commenting on the curious word known only from *CIL* 1 ed. 2, 1614 and Buck 40² (Pisani 31B). The correct reading is *fancla*, cf. Varro LL 5.137 *fancla* (HVp: *phancla* fG), to be compared with Sicel ζάγκλον (*PID* 2.580) and Ligurian (?) *daculum* *CGL* 1.84, 91, Latin *falx* 'sickle', here sens. obsc. The curse invoked is priapism, a punishment to fit the presumed crime.

Der Aufbau der Sprache. By BRUNO SNELL. Pp. 221. Hamburg: Claassen Verlag, 1952.

Reviewed by WERNER WINTER, *University of Hamburg*

The author, a well-known classical scholar with a keen interest in the philosophy of language and in general linguistics, tries to show that the whole structure of a language is determined by Bühler's three categories of linguistic function: *Auslösung*, *Kundgabe*, *Darstellung*, which Snell calls 'Wirkungs-, Ausdrucks-, Darstellungsfunktion'. To prove his thesis, he goes even farther than Junker—Streitberg *Festschrift* 14 (1924)—in coordinating forms and grammatical cate-

gories with these basic functions. The book as a whole is highly interesting, less for new insights in problems of detail than for its pervading attitude. Throughout the book one feels the author's conviction that language, as a product and reflection of mental processes, can be adequately understood only by the methods of *Geisteswissenschaft*. Thus, the book is to be taken primarily as an attempt to develop and apply such a method in the field of linguistics.

A linguist who is accustomed to the positivist approach and therefore tends to suspect all philosophical, logistic, or artistic approaches to language will be shocked by many of the conclusions in Snell's book. He will probably object first of all to the author's limited materials; for Snell confines himself to the languages he knows best: Greek, Latin, and some languages of modern Europe. The reviewer does not feel that such a limitation is necessarily a shortcoming; it seems, rather, a necessary safeguard in performing the extremely subtle analysis on which to base conclusions about the mental categories reflected by those of grammar. It would, however, have been better if the author had claimed no more for the results of his investigation than that they apply to the languages actually discussed, without risking conclusions about language in general. As a title, *Der Aufbau UNSERER Sprache* would have been more realistic than the one chosen. In defining the elements which he finds in the languages discussed as 'Urphänomene', the author tacitly assumes something like a universal parallelism, or even uniformity, in human thought; but a study of more exotic languages would have shown him clearly that patterns of thought are radically diverse throughout the world.

The author's coordination of grammatical categories with the basic functions is not always equally persuasive. Since he does not go so far as to ascribe to each category only a single basic function, but rather sees several at the same time with one or another predominant, his interpretations often seem highly subjective; but probably this is unavoidable as long as no method generally agreed upon is available. By the same token, objections to these interpretations will also often be just as subjective, and hence not always more convincing. Certain points in the author's discussion, however, call forth objections for more tangible reasons. A few of these points will be mentioned here.

The author ascribes to the phonemes (41) a primarily expressive value—that is, he correlates them with the second of Bühler's functions. From the highly expressive character of some interjections, he infers (48) that in certain words the sounds have expressive values identical with those of the interjections in which they also occur. Now it is well known that attempts have been made, time and again, to abstract the meaning of a sound from certain words in which it occurs, and to use the apparent results of the analysis to detect hidden layers in poetical and other texts. Though some of these attempts do not lack a certain persuasiveness (the examples given by Snell are among the best), it can easily be shown that the apparent connotations of sounds are valid only, if at all, for one language at one particular time (cf. Snell's examples, 50). One would therefore have preferred the author not to try a transfer of the rather vague results obtained for present-day German to languages like Latin and Greek (51, 53). It is hard to reconcile the scientist's statement (50) that sound change is contradic-

tory to an overall system of sound symbolism, and the artist's attempt to ascribe a certain meaning to some sounds (50 ff.): the reader is left with the uneasy feeling that the tendency to speculate is maintained in spite of the author's awareness of its inadequacy.

Determining the value of sounds from their use in interjections seems at first glance much more reasonable; but even here one must object to the conclusions offered. Since the sounds of interjections often fall outside the normal phonemic pattern of a language, even those sounds of interjections that seem to be allophones of ordinary phonemes had better not be thus identified. Compare C. M. Doke's distinction between normal grammatical phonetics and extranormal phonetics, *A study in Lamba phonetics, Bantu studies* 3.6 (1927).

While sounds are said to be primarily expressive, words are correlated with the category of *Darstellung* (60 f.) and sentences with that of *Wirkung* (66). The latter attribution is partly based on the fact that the verb (which the author finds to be correlated with the category of *Wirkung* too, cf. 94) is regarded as the dominant factor in the sentence. Adjectives and nouns are related to the expressive and the descriptive functions respectively. In a similar way virtually all grammatical categories are fitted into the scheme. Sometimes it is obviously impossible to avoid a rather arbitrary distribution; for instance, it is hard to see why imperative (110 f.) and aorist stems should be related to this same category of *Wirkung*, until one discovers that in the first case *Wirkung* is taken to mean 'effect', while in the second one it means 'action'.

Such criticisms of details could easily be continued. But one should not overlook the merits of Snell's enterprise. Even if one cannot accept his conclusions as following strictly from the linguistic evidence, one may welcome his attempt to achieve an overall picture of his language as a reflection of thought, feeling, and will. It is granted that such an attempt cannot succeed without the detailed work of the linguists; but the latter must see, for their part, that their studies will fall apart into unrelated specialties unless they develop a philosophy to which they can relate their work and its results.

Die Entwicklung neuer germanischer Kultursprachen von 1800 bis 1950.

By HEINZ KLOSS. (Schriftenreihe des Goethe-Instituts, Vol. 1.) Pp. 254. München: Pohl & Co., 1952.

Reviewed by WERNER WINTER, *University of Hamburg*

This is a book designed more for a public of interested laymen than for students of linguistics. But since it contains valuable information hard to find elsewhere, it seems, in spite of its obvious shortcomings, to deserve a review in this journal.

Kloss explicitly states that his interests and intentions are those not of a linguist, but of a sociologist of language. What he wants is to describe and analyze the development of a given idiom from the status of 'dialect' (in the popular, somewhat condescending sense) to that of an all-round 'language of civilization'. He tries to establish criteria for determining the stage reached by a particular idiom at a particular time, and to place various idioms in the system thus developed.

The chief merit of the book, as already stated, lies in its collection of data about a number of imperfectly known Germanic idioms, such as Saramacca and Beach-la-mar (accepting for the moment the author's classification of these languages as Germanic), and in its outline of recent developments in others better known, such as Low German and Swiss German. The term 'development', of course, is to be taken in the sense mentioned above: toward the status of a full-fledged *Schriftsprache*. The author's bibliography cannot be anything but selective, particularly for the better known idioms; no one will blame him for not giving a full account of all the available information.

There are certain points in which the author's information seems to have been insufficient. Thus, the statement that the local Swiss German idiom is not used in church except in a few remote valleys is not correct: it has become quite customary to conduct evening services in the local 'dialect' (I have recently witnessed this in the city of Bern and in eastern St. Gallen); only the more formal Sunday morning service seems to be the domain of *Schriftdeutsch*. I can also testify to the use of the local idiom in the first grades at least of rural schools as late as 1949-50. Furthermore, Kloss should not have overlooked the part played by this idiom in daily newspapers.

In the theoretical sections of the book we find some good observations about the popular use of the terms 'dialect' and 'language' (15 ff.). It is quite true that the social reputation of an idiom classifies it decisively in one category or the other. This goes even farther than the author thinks: the 'linguistic uniqueness' which he lists as the chief criterion for conferring the rank of 'language' upon an idiom plays a much smaller part than the irrational social point of view. Thus it is significant that again and again we find expressions like 'Negro dialects' and 'Indonesian dialects', relics from a time when it was felt that 'such people' could not possibly have 'languages'. Although this irrational element is very strong in the popular use of these terms, the author of a sociological study need not abandon the purely social classification altogether; but at least he ought not to mix popular terms with linguistic terms that are unfortunately homonymous. In a strict sense, the linguistic terms 'language' and 'dialect' express only varying degrees of similarity between two idioms. To ask whether the idiom X, taken by itself, 'is' a language or a dialect makes no sense; X can be called one or the other only in comparison with some second idiom Y. Even if Y is canonized as a *lingua franca* throughout an area that includes the area of X, this X cannot be labeled a dialect of Y unless the name Y is used as a cover term for a whole group of closely related, basically equal idioms including both X and Y. Furthermore, if X is compared with a wholly dissimilar idiom Z, it is not correct to call X a dialect: compared with Z, X is another language, just as Y is another language. In such a framework there is no room for an evolutionary and evaluative classification like the one used by Kloss. If one wants to pursue a study of this kind, it would seem better to refrain entirely from using the technical linguistic approach side-by-side with the popular social one.

Among the author's valuable contributions is his statement that the degree of development of a *Schriftsprache* is determined by the development of non-fictional prose. But such contributions are unfortunately often spoiled by his

rather dilettante elaborations concerning problems of linguistic theory. Kloss tries to deny the primacy of linguistic structure over vocabulary by ascribing structure to the nationality of the speaker and then assuming that what he considers the true intralinguistic parts of language—vocabulary, phonology, and morphology—are the only important ones. But such wanderings in an area where the author is obviously not at home scarcely impair the value of this book as a source of facts. As for the theories, one can accept only a part of them, and that part only within a non-linguistic frame.

One last criticism: the author's terminology is often poorly chosen. What is the use of a transcription *Bietschlar* for *Beach-la-mar* (of which he writes, 'Zur Beibehaltung der schwerfälligen engl. Schreibweise besteht kein Anlass!'), or of the term *Pennsilfaanisch*? And words like *umvolken* or *Angelamerikaner* just sound odd.

Reflexive und intransitive Verba im älteren Westgermanischen. By LARS HERMODSSON. Pp. 347. Uppsala: Almqvist & Wiksells Boktryckeri AB, 1952.

Reviewed by ALBERT MOREY STURTEVANT, *University of Kansas*

This monograph presents a complete picture of the syntactical development of reflexive and intransitive verbs in the early stages of the West Germanic dialects. The author has treated his subject from a comprehensive point of view, devoting an entire chapter to Gothic, with incidental references to Old Norse wherever the two dialects show a syntactical affinity in contradistinction to West Germanic. He has spared no pains in documenting his evidence. His conclusions are based strictly upon this evidence, and whenever the evidence is not convincing they are stated with scholarly caution. The whole work gives the impression of orderly arrangement, accuracy in details, clarity, originality, sound judgment, and enormous industry. The author's main contribution is that he has here woven together the syntactical characteristics of these verbal types into one organic whole. Through this coordination Hermodsson substantiates his contention that some of the traditional theories regarding Germanic verbal syntax must be discarded, notably when the traditional viewpoint cannot satisfactorily explain certain syntactical phenomena in the West Germanic dialects. To provide a satisfactory method of exposition Hermodsson takes his point of departure from the Indo-European syntactical antecedents, tracing their survival in Germanic under the conditions imposed by secondary Germanic peculiarities—a coordination of historical evidence.

The work is divided into eleven chapters: 1. Allgemeine Einleitung; 2. Zum Gotischen; 3. Zum Alt- und Mittelhochdeutschen; 4. Allgemeines zum Reflexivum und Intransitivum im Nordwestgermanischen; 5. Zum Altsächsischen; 6. Zum Altenglischen; 7. Zum Altfriesischen; 8. Zum Mittelniederländischen; 9. Exkurs: Das Eindringen des reflexiven Pronomen *sich* ins Niederländische; 10. Zum Mittelniederdeutschen; 11. Zusammenfassung, Zitierte Literatur, Alphabetisches Verzeichnis der behandelten Verba.

The introductory chapter (13–82) gives a history of the terminology connected

with verbal syntax; an analysis of the verbal diatheses as determining the correctness of this terminology, with special reference to the passive and middle or medial usage; a discussion of the form of the reflexive pronoun and its syntactical relation to the verb, and of personal and non-personal reflexive verbs; a history of the terminology and definitions of the intransitive verb; an analysis of the syntactical nature of the intransitive verb; a discussion of verbs with double function, i.e. verbs used both transitively and intransitively. The introductory survey thus furnishes a basis for the author's method of procedure in applying his principles concerning syntax to the West Germanic languages. This naturally involves much repetition in the following chapters where his principles are applied, but the repetitions are justified so far as they clarify and coordinate the argument, as is attested by the indispensable cross references. Because of the vast scope of the survey I must confine myself to a discussion of a few salient features.

In regard to the *nan*-verbs in Gothic, Hermodsson rightly regards these verbs not as genuine passives but as 'mediale Verba' (59). He does not tell us, however, that the *nan*-verbs have an inchoative sense (cf. Albert E. Egge, Inchoative or *n*-verbs in Gothic, *AJP* 7.38-45), nor does he mention the fact that they translate the Greek mediopassive only when this has a medial sense (cf. Braune, *Gotische Grammatik*¹⁰ §194). The difference between the diathesis of the medial and the passive voice he explains as follows (23): 'Der Unterschied zwischen einem Vorgangsausdruck wie *die Netze zerrissen*, got. *natja dishnupnodedun* ... und einer Passivform wie *die Netze wurden zerrissen* erkennt man einfach daran, dass die letztere die Bedeutung des entsprechenden aktivischen Ausdrucks hat, "(jemand) zerriss die Netze", während die erstere Wendung nicht auf diese Weise wiederzugeben ist.' This statement is of course true: the verb *dishnupnodedun* has an intransitive-medial sense since it denotes the act of 'tearing apart' without reference to any agent responsible for the act. But Hermodsson fails to state the fact that this agent (person or thing) may be IMPLIED in the context; cf. Luke 5.6 *jah þata taujandans galukun manageins fiske filu, swe natja dishnupnodedun ize* 'after they had done this, they caught (enclosed) a great multitude of fishes, so that their nets were torn apart'. In this passage it is clear that the reason why 'the nets got torn' (inchoative) was that there were so many fish in them; or in other words, 'the fish' represent the implied but unexpressed agent which caused the nets to get into this condition. It is just at this point, where medial and passive blend, that the two syntactical functions have not been clearly differentiated by those who construe the *nan*-verbs as passive. Hermodsson does not make clear this point of contact by a mere reference to it (29): 'In Ausdrücken mit der Vorgangsbedeutung ist das Subjekt nicht mehr das regierende Glied, sondern geradezu dem Verbum untergeordnet. In diesem besonderen Zug liegt eine gewisse Übereinstimmung mit dem Passivum.'

In regard to the restriction of the passive usage, Hermodsson says (24): 'Im Passiv werden meistens gerade die Verba mit Tätigkeitsbedeutung gebraucht. Verba wie *erhalten*, *bekommen* werden nicht passivisch verwendet.' Just exactly what is meant by 'Verba wie *erhalten*, *bekommen*' is not clear. Both these verbs have various senses, and both *erhalten* 'to receive' and *bekommen* 'to get, receive'

may be used in the passive voice, e.g. *Jeden Tag werden vom Bibliothekar neue Bücher erhalten* (bekommen).

Very important is the discussion of the reflexive pronoun IE **se-* **sye-* and its congeners in the Germanic languages. As Hermodsson here points out (30), this pronoun was originally used in both an anastrophic and a reflexive function, but gradually the reflexive function superseded the others. Relics of this double usage are found in both Greek and Latin (31-2): in the Greek anastrophic *αὐτόν* for the reflexive *ἐαυτόν* (*ἐ-* < **se-*), and in the Latin *se* of a subordinate clause, referring to the subject of the principal clause, e.g. *Caesar omnibus qui contra se arma tulerunt, ignovit*. This original double function of the reflexive form survived in the Germanic languages: the Latin usage especially in Gothic and Old Norse (183), and the Greek usage especially in the 'ingwäonisch' dialects of West Germanic, where the personal pronoun was substituted for the reflexive and finally supplanted it (183-5). The author's hypothesis seems well founded on historical evidence, and thus represents an advance in research; for although scholars have noted the anomalous use of the reflexive pronoun in Germanic and the substitution of the personal anastrophic form (cf. Eng. *him*, *himself*, etc.) for the reflexive, they have never given a convincing reason for these syntactical and formal divergences.

The discussion of the verbs with double function (60-82) has an important bearing upon the assumed omission of the reflexive pronoun as object of a transitive verb. Here the author points out that Proto-Indo-European must already have had intransitive verbs in which the medial diathesis could be expressed by the active form, so that the usual active form could express now the active and now the medial sense. According to Hermodsson this syntactical status survived in Germanic, especially in denominative *jan*-verbs, which were used with double function from the time that they were first formed. Such verbs the author calls 'detransitiv', a term which differentiates them from those verbs which were primarily only intransitive. Detransitive verbs never had a reflexive object, and therefore the usual explanation that an object is omitted is erroneous; the so-called 'absolute' usage of transitive verbs in an intransitive function represents a survival of the original IE status, in which the active form could express the medial (intransitive) sense. The extension of the double function was favored by the congruence of the transitive and intransitive verbal forms; cf. NHG *verderben* 'to spoil, corrupt' < OHG *farderbian* trans. and *fardērban* intr., MHG *stillen* 'to quiet' < OHG *stillōn* or *stilljan* trans. and *stillēn* intrans.

In the following chapters, the syntactical principles laid down in the introduction are applied to the Germanic languages. Each language is investigated according to uniform categories: A. Reflexive Verba, B. Verba mit Doppelfunktion. Under A there is a subcategory, Nicht persönliche Reflexiva, included only in the chapters devoted to Gothic and to Old and Middle High German, where the nonpersonal reflexive verb plays an especially important role. In the following discussion I shall confine myself to details, since the general principles involved have already been discussed above.

In Chapter 2 (83-108), the author's chief task is to determine the native Gothic syntactical usage. His method is to prove that the Gothic syntactical usage ex-

isted in the other Germanic languages if foreign influence cannot be assumed, thus eliminating the possibility of Greek or Latin influence upon the Gothic construction in question. The soundness of this method is illustrated when the Gothic verb apparently expresses a different diathesis from that of the Greek verb which it translates. The Gothic denominative intransitive *jan*-verbs usually denote a condition, not a medial inchoative sense; cf. *riqizjan* 'to BE dark'. The verb *riqizjan*, however, occurs in Mark 13.24 (*riqizeip*) as a translation of the Greek inchoative medial σκοτισθήσεται 'shall be darkened (grow dark)'. That the inchoative sense 'to GROW dark' of the verb *riqizjan* in this passage was inherent in the Gothic *jan*-verb and therefore independent of Greek influence is proved by the existence of the OE inchoative *jan*-verbs, such as *hliewan* 'to grow warm', *drýgan* 'to become dry' (95).

Hermodsson's hypothesis (98) that Gothic *augjan* 'to show', instead of being a denominative verb derived from the substantive *augo* 'eye', is based upon the same Gmc. root **aug-*, seems to me fully justified, since otherwise—though Hermodsson does not mention this in support of his hypothesis—we might expect the verb *augjan* to signify 'to see'; cf. Gk. ὀφθαλμός 'eye': ὀψομαι 'I shall see', Lat. *oc-ulus* 'eye': *oc-ulāre* 'to furnish with eyes, to make see', and Eng. *to eye* 'to fix the eyes upon, to view, observe'.

'*Anastodjan* ist offenbar als Kausativum, d.h. mit trans. Funktion gebildet worden' (104). This statement is misleading insofar as it implies that the transitive function here developed out of the causative function. The causative verb corresponding to the intransitive *standan* 'to stand' is *gastōþ[an]an* = στήσαι (Rom. 14.4), which is evidently a secondary formation derived from an adjective **gastōþs*; cf. *un-gastōþs** 'not having a STANDING (permanent) abode' (*ungastōþai* = ἀ-στατ-ούμεν, Cor. 4.11). Hermodsson has here overlooked the fact that the IE *eje*-suffix, from which the Gothic *j*-suffix is derived, could denote not only a causative but also a frequentative or iterative force (cf. Brugmann, KVG §698.1). According to Jellinek (*Gesch. d. got. Sprache* 196 n.), this iterative force was originally present in Gothic *-stōdjan*, but had so faded that the sense of the secondary form with *j*-suffix became identical with that of the primary verb *standan*, as represented by OHG [*gi-*]*stantan* 'to begin'. This explanation satisfactorily accounts for the non-causative sense of Gothic *anastōdjan* 'to begin', especially since there are no examples in the Gmc. languages of the same verbal form in causative and non-causative function. For a discussion of this question see my article, Gothic miscellanea, *Germ. rev.* 28.58–9 (1953).

In regard to the omission of the reflexive pronoun *sik*, the author points out that in certain examples this may be due to the supplanting of the original native sense of the verb by a new meaning, modeled on the foreign verb which the Gothic translates; thus *daupjan* 'to wash': *daupjan sik* 'to wash oneself' > *daupjan* = βαπτίζειν 'to baptize', a technical term connected with ecclesiastical rites (cf. 44). In most examples, however, he believes that the omission of *sik* cannot plausibly be laid to the influence of a corresponding Greek or Latin verb; an example is *bimaitan* 'to cut': περιτέμνειν 'to circumcise'. 'To cut oneself (voluntarily)' or 'to have oneself cut' could hardly have been a native conception common enough to succeed in establishing the use of the reflexive *sik* with

bimaitan, parallel to *daupjan sik* 'to wash oneself', a common act of everyday life (105). Hermodsson explains the intransitive *bimaitan* as a detransitive verb.

The chapter presents a new interpretation of Gothic syntax. According to Hermodsson, the syntactical structure of the Gothic Bible on the whole reflects native usage to a far greater degree than Streitberg and his adherents were willing to admit. Hermodsson's conclusions seem to confirm Curme's contention to this effect (Is the Gothic Bible Gothic?, *JEGP* 10.151-90, 335-77 [1911]); but whereas Curme relied primarily on Sprachgefühl, Hermodsson bases his conclusions on comparative and historical evidence. Whether we agree with his hypothesis regarding detransitive verbs or not, it is a new viewpoint which deserves serious consideration.

In Chapter 3 (109-77), Old and Middle High German are treated together because they show no fundamental difference in verbal syntax. The OHG material is based upon Tatian, Otfrid, Notker, and Williram; the MHG upon the chief literary monuments up to about the year 1200. Hermodsson shows (118 ff.) that the percentage of reflexive verbs with nonpersonal subject was much greater in OHG than in Gothic, but was never exceeded in MHG. The latter fact he explains on the ground that as MHG writers became more concerned with abstract notions, things (nonpersonal subjects) were conceived as living beings (personal subjects) and were more frequently personified, especially in poetry. As for the reflexive usage of transitive verbs with a personal subject, Hermodsson does not believe that the Latin influence of the medial *r*-forms was responsible for this usage to the degree that is generally assumed (139). He illustrates the unreliability of explaining *frewen sich* by *laetārī* and *mīdan sich* by *verērī*, when *frewen sich* just as often translates *exultāre* and *mīdan sich* usually translates *ērubēscere* and never *verērī*. The assumption of French influence is entirely vitiated by the fact that the French reflexive usage did not develop until the 14th century (141). Aside from the few examples in which Latin influence may be assumed, the reflexive verb in Old and Middle High German (as in Gothic) clearly represents native Germanic usage. The following syntactical and formal elements were involved in the competition between transitive and intransitive function and thus determined the extent to which the double function developed in OHG (142 ff.): the *jan*-verbs, most often transitive; the *ōn*-verbs, either transitive or intransitive; and the *ēn*-verbs, intransitive as cognate counterparts of the transitive *jan*-verbs (cf. *nazēn* 'to be wet' : *nezzen* 'to make wet').

The MHG use of an intransitive verb when the corresponding transitive denotes 'motion away from' (as *wenden* 'to turn', *trīben* 'to drive', *sprengen* 'to make run') is traditionally explained by ellipsis of a direct object (such as *ros* 'horse' or *schif* 'ship'). Hermodsson discards this theory as not warranted by the evidence (156): 'Es wird dem Pferd hier wirklich eine recht grosse linguistische Rolle zugeschrieben' (159). He notes that there are a few examples where *ros* is used as the object of such verbs, and admits the ellipsis could conceivably have occurred; but the examples are too few to account for the generalization of this type of intransitive verb. Rather, it can be satisfactorily explained as detransitive (156); i.e. the transitive verb underwent a change in meaning which resulted in a medial diathesis (intransitive). The same principle is applied to

intransitive verbs of the type *rüeren* 'to move', alongside *sich rüeren*: the transitive reflexive denotes 'motion in a fixed position' ('Bewegung an Ort und Stelle'), whereas the intransitive denotes 'motion away from' ('Fortbewegung') and has thus acquired a new sense in its detransitive function (160).

Chapter 4 (178–85) sketches the overall picture of verbal syntax in Northwest Germanic ('ingwäonisch') as contrasted with that of OHG–Gothic, with particular reference to the reflexive pronoun. Hermodsson points out that IE **se-*, **sye-* was used in the double function of reflexive and anastrophic pronoun, with the result that in NWGmc. the reflexive form survived only in the accusative case (just as in OHG), the dative form being supplanted by the personal anastrophic form (180–81). One important circumstance which favored the loss of the reflexive form in NWGmc., over against its retention in OHG, is the fact that reflexive verbs were used much less frequently in NWGmc. than in OHG; indeed, certain verbal categories of OHG are either entirely or partly lacking in Northwest Germanic. (184).

Chapter 5 (186–92), covering the Heliand and the Genesis-fragments, is rather limited in scope, but nevertheless reveals the characteristic features of NWGmc. syntax. One of the distinctive features of OS verbal syntax is the frequent use of the dative anastrophic pronoun in a reflexive function with intransitive verbs, as compared with the extremely rare use (with the reflexive form) in OHG; cf. *he geng im tho bi Iordanes staðe* (Heliand 1127). Hermodsson notes, however, that this dative does not bear the same syntactical relation to the verb as does a pronoun used in the reflexive function as indirect object of a transitive verb; it is obviously a *dativus commodi* or *dativus ethicus*. Here the author might well have pointed out the parallel usage in the modern Scandinavian languages. Here it was introduced from Low German and restricted to poetry (especially the folk ballads), which rouses the suspicion that the frequent use of this dative with intransitive verbs in Old Saxon was due to the poetic style of the Heliand and Genesis; cf. Norw. *Jeg gikk mig en gang* : OS *he geng im tho* ...

Chapter 6 (193–210) treats the Old English material, which is far more extensive than the Old Saxon and exhibits the 'ingwäonisch' features far more clearly. Detransitive verbs, characteristic of 'ingwäonisch', are more numerous than in OHG; cf. OE *hebban* 'to lift' trans. and intr. : OHG *heffen* trans. and refl. (209). The author notes that the 'detransitivizing' process has progressed further in Modern English than in any other Germanic dialect; he assumes that this is due chiefly to the peculiar phonetic development of Modern English: 'offenbar vor allem infolge der besonderen lautlichen Entwicklung des Ne.' (210). It is not clear how he arrives at this conclusion inasmuch as he does not explain the relation of form to syntax which brought about this result.

In Chapter 7 (211–6), the Old Frisian material is based almost exclusively upon legal documents. The main characteristics of 'ingwäonisch' syntax are clearly preserved, but the dearth of material forces the author to use great caution in drawing conclusions from the evidence. Some of the OHG categories are entirely lacking, and the reflexive verbs play a less important role than in Old High German.

Chapter 8 (217–62) investigates material of very extensive scope. The status of

Middle Low Franconian is obscured by the intrusion of Middle Low German elements; and the comparison with OHG and MLG data is complicated by the fact that these are drawn chiefly from literature written before 1200, whereas the MLFr. data are scattered between the 13th and 16th centuries. Nevertheless, the author succeeds fairly well in giving us a synchronous picture, in which all the 'ingwäonisch' syntactical characteristics are clearly brought out. He discusses in detail the contacts between Franconian and Middle Low German, due to geographical proximity and political influences. One of the most baffling problems is the intrusion of the MLG reflexive pronoun *sik* (*zich, sich*); this has such an important bearing on the status of MLFr. reflexive verbs that the author devotes Chapter 9 (263-91) exclusively to this topic.

Agathe Lasch, one of the authorities on Middle Low German, considers the MLG form *sik* to be of native origin (*Mnd. Gr.* §337). Hermodsson disproves this assumption by showing that the High German form *sich* turns up in Middle Low German also, whence it is highly probable that the pronoun was borrowed by Middle Low German from High German. The MLG form *sik* is then a re-forming of HG *sich* on the pattern of *mik* and *dik* (293-4). The native MLFr. usage of the reflexive verb is represented by the anastrophic form (*hem*) as object of the verb, whereas the parallel usage with the reflexive form (*zich, sik, etc.*) represents MLG influence. These two combinations the author designates as 'hem-Verbindungen' and 'sich-Verbindungen'. He discusses in detail the competition between them in the MLFr. dialects, and concludes that the intrusion of MLG *sik* is not merely the exchange of one morpheme (*hem*) for another (*sik*), but brought about a fundamental syntactical trend in favor of the reflexive construction, displacing the native MLFr. detransitive pattern (290).

In Chapter 10 (292-303), Middle Low German is not classified with Old Saxon because the former shares the syntactical characteristics of Middle High German. The author has selected most of his material from MLG dictionaries, a procedure which does not yield such satisfactory results as examination of texts. To ascertain the native status of MLG syntax is as difficult as in Franconian: not only the MLFr. forms but also the Frisian and the MHG must be eliminated. Hermodsson carefully works out this problem, especially in connection with the rivalry between 'hem-' and 'sik-Verbindungen'. He concludes that the 'sik-Verbindung' in the dative reflexive construction was due to MHG influence and thus helped to preserve the old dative usage which had disappeared in 'ingwäonisch'—to which MLG, aside from its syntax, properly belongs (298). Verbs with double function occur with approximately the same frequency (against those with transitive and reflexive uses) as in MHG (301). The chapter shows clearly how a LG dialect merged syntactically into the HG group.

Chapter 11 (304-47) deserves special comment because it contains everything necessary for a retrospective evaluation of the monograph: a lengthy summary of the author's treatment, with numerous cross references; a complete bibliography of the texts upon which the investigation is based, of the dictionaries and journals consulted, and of the philological literature; and finally an index of the verbs discussed. The book is beautifully printed; I have discovered only one disturbing misprint, *tings* for *things* (200, under *yfelian* 1).

Hermodsson has produced a masterful treatise; nothing is neglected. Working objectively, he has given us an essentially new evaluation of the whole problem, which may cause us to revise some of our orthodox theories regarding verbal syntax. His view of the syntax of reflexive and intransitive verbs in Early West Germanic is so convincingly substantiated by Gothic and Old Norse parallels, and so well founded on the IE evidence, that no linguist can refute it without a new investigation. Even if this is undertaken, it is doubtful that Hermodsson's fundamental principles, contradicting so much of the traditional speculation upon this subject, can ever be disproved.

Middle English dictionary. By HANS KURATH, editor, and SHERMAN M. KUHN, associate editor. Part E.1, *ē* to *ēndelþnges*, pp. ii, 120. Ann Arbor: University of Michigan Press, 1952.

Reviewed by KEMP MALONE, *Johns Hopkins University*

This dictionary is one of a group of large lexicographical undertakings set going in the English field as the NED or 'big Oxford' neared completion. The first of these undertakings, the *Dictionary of American English*, came out in four volumes during the years 1938-44, under the auspices of the University of Chicago. A *Dictionary of the older Scottish tongue* (to 1700) is also under way at that University; publication began in 1931 but is still far from complete. The University of Michigan undertook to compile a dictionary of early modern English; but this enterprise has been shelved until the completion of the *Middle English dictionary*. The latter undertaking, originally launched at Cornell University with Clark S. Northup as editor, was transferred in 1930 to the University of Michigan, where it has remained. The first editor after its transfer was Samuel Moore. After his untimely death in 1935 the editorship passed to Thomas A. Knott; and after Knott's death in 1945 it passed finally to Hans Kurath, who has now brought the work to the point where publication can begin.

The present fascicle, which came out late in 1952, is the first to appear. In its Prefatory Note we are told that 'for special reasons, E and F will be published first. They will be followed by D, C, B, A, in that order, whereupon G and the remaining letters will appear in alphabetic sequence.' The publishers announce, on the back cover, that 'the completed *Dictionary* will run to approximately 8000 pages; it will be issued in 124-page parts, five or six a year, over a period of ten years.' But from the Prefatory Note we learn that only three parts will come out in 1953, one of them devoted to 'a full description of the editing plan, the bibliography, ... a detailed acknowledgment of the contributions of former members of the Dictionary staff,' etc. In the present fascicle appear 'only the briefest comments on the editing plan' and the reviewer will have to be even briefer in dealing with this matter. It will be enough to quote the following from the Prefatory Note.

The entry forms are based upon an analysis of the systems of sounds and forms current in the Southeast Midland (or London) dialect of c1400. The spellings chosen for the sounds conform rather closely to the spellings of the Chaucer manuscripts, except that the vowels *i* and *ī*, and the diphthongs *ai/ei*, *oi/ui*,

eu, au, ou are always written thus (never *y, ay/ey*, etc.), so that the derivatives may directly follow the base form, e.g. *dai, daiġ; hōlġ, hōlġnesse*. The spelling of the entry forms is not strictly phonemic. (a) Whenever one and the same sound is more or less consistently spelled in more than one way, the historical spelling of the particular word is used in the entry form: e.g. *dai*, but *wei* ... (b) Although the long intervocalic [sic] consonants of Old English lost their phonemic function in early Middle English as the result of the lengthening of short vowels in the open syllable, and henceforth varied automatically in length (being short after long stressed vowels and diphthongs, and long after short stressed vowels, as in Modern English), double consonants are written in the entry forms in accordance with the historical spelling: e.g. *fillen, sitten, likken* ... The phonemic value of ambiguous spellings is identified by diacritics ... If a word is not attested in the Southeast Midland dialect, or if it occurs only in Early Middle English, the entry form is based upon the phonology, or the spelling, of one of the other dialects—Northern, Western, or Kentish—or upon the spelling of one of the early texts in which the word occurs. All quotations for one and the same word ... are, in principle, treated under one entry form. This procedure makes possible a full analysis of the range of meanings of an expression throughout the Middle English period (1100–1475) in all parts of England.

This plan is obviously practical rather than rigorous or doctrinaire. I will comment only on item (b), which deals with the double consonants of Middle and Modern English. It is wrong to say that in Modern English intervocalic consonants are 'long after short stressed vowels'. Length in this position, though common in Middle English, is rare in Modern English. Both words of the riming pair *rabbit/habit* have short [b] and the doubling of the consonant symbol, where it occurs, is only orthographical. One finds a truly double consonant in phrases like *mad dog*, where the first *d* ends, the second *d* begins a morpheme; but apart from such cases (which do not depend on the length of the preceding vowel) long intervocalic consonants are foreign to current speech. Again, it is wrong to say that the long intervocalic consonants 'lost their phonemic function in Middle English'. In oppositions like *son* 'son': *sonne* 'sun' and *spite* 'tool for cooking meat': *spitte* 'eject saliva' the long consonants are clearly phonemic. And it will not do to say that intervocalic consonants 'varied automatically in length, being short after long stressed vowels and diphthongs and long after short stressed vowels'. The opposition *write* 'write': *wrote* 'wrote' (pl.) shows that an intervocalic consonant might be short irrespective of the quantity of the preceding vowel. Luckily these errors of interpretation were not carried over into the body of the dictionary, where 'double consonants are written in the entry forms in accordance with the historical spelling'.

A given entry in the body of the dictionary starts with the 'entry form' (see above) and the grammatical classification of the word. Then come variant forms, followed by the etymology in square brackets. These items make the first section of the entry. The rest consists of definitions and supporting quotations. If there is only one group of definitions, this comes before any of the quotations, the individual definitions being marked (a), (b), etc. and the quotations following in sequences likewise so marked. If the definitions fall into two or more groups, each group is given a number and the quotations that go with it come before the next numbered group of definitions. The quotations are identified by source and date; the latter is the date of the MS that contains the text quoted from, but if

the date of composition of the text differs appreciably from that of the MS, both dates are given. Precise dating is often impossible, of course; in particular, the date of composition is given simply as OE if the text quoted from, though 'modernized' (i.e. Middle English) in linguistic form, goes back to OE times. Quotations from official records, and from a few major writers (e.g. Chaucer), may be dated by time of composition only.

It will be worth our while to examine one of the entries somewhat closely, in the light of the procedures just explained. For this I have chosen the entry *eit* 'island'. The entry form *eit* belongs to the West Midland dialect. A variant form *ait* is also given, although this occurs in none of the quotations. On the other hand, none of the variants that do occur in the quotations are listed with *eit* and *ait* in the first line of the entry. Clearly the variants set alongside the entry form need not be more representative of the records than is the entry form itself. The form *eit* is said to be 'from OMerc. **ēgob*', but this connection leaves the *t* unexplained. A better hypothetical form would be **ēget*, answering to the recorded WS *ȝgett*. The suffix *-et(t)* here is to be identified with that in OE *emnet* 'plain', *piccett* 'thicket', etc. The *portmaneit* in the quotation from the Godstow Cartulary reads *portman eit* according to EETS, OS 129.29 n. 1. In the same volume, 322 line 23, occurs the form *heyget* (the *h* is presumably inorganic). Another variant *neit* is recorded in 130.653 line 22. Both these variants are of interest, and it is a pity they did not win record in the MED. Under *eit* the editors also take up the compound *eitlond* 'island', and this arrangement has its advantages, but a cross reference might well have been given for the benefit of those who look for *eitlond* in its alphabetical place.

One may suspect that the variant *ait* was singled out, in spite of its absence from the quotations, because it is the usual modern (dialectal) form. The same may be said of *echo*, given as a variant of *ecco* though not found in any of the quotations. The choice of an entry form makes special difficulties in the large class of words ending in *-ion* or *-ioun*. One might expect to find *-ioun* for words taken from French, *-ion* for words taken from Latin; or the entry form might be made to agree with the form oftenest found in the quotations, if one could not decide between French and Latin origin. If the word is said to be from Latin, and if no forms in *-ioun* appear in the quotations, one would have a right to expect *-ion* in the entry form, and this is actually the case with *edicion*; but we find *ejeccioun* as entry form in spite of the etymology given and the evidence of the quotations. Presumably the editors did not regard system or consistency in this matter as having much importance and preferred to spend their time on things of greater moment.

The entry '*ek* adv. and conj.' gives us more material: 53 occurrences of the word in 52 quotations. The form *ek* chosen to head the entry occurs seven times in the body of quotations. Of the four forms listed in the first line of the entry as variants, *eac* occurs in the quotations 10 times, *iec* and *yke* once each, and *zeke* not at all. Of the eight variants that occur in the quotations but are not listed in the first entry line, *eke* occurs no less than 20 times. Next in frequency come *ec* with five and *eek* with four occurrences. The other five, *ech*, *heke*, *eek*, *æc*, and *æc*, occur once each.

No doubt *ēk* heads the entry because it agrees with the normalized spelling used in this dictionary for the entry forms. The variant *eac* was listed in the first line, one may suppose, to represent early ME orthography. The other three variants listed as such in the first line are forms of special interest to phonologists and dialectologists. Yet it is odd to find *eke* left out of this list. The form agrees with *yke* (and with the unlisted *heke* and *eeke*) in having an added *-e*, presumably the adverbial suffix. This feature gives *eke* a special interest. But frequency of occurrence alone is enough to justify if not indeed to require the inclusion of *eke* among the variants listed in the first line of the entry.

The editors content themselves (rightly, no doubt) with diacritical marks to supplement the conventional spelling in cases where this is ambiguous. In other words, this dictionary gives no respellings in phonetic script to indicate pronunciation. In the interests of simplicity, moreover, words that vary in pronunciation (compare modern English *either*) may be treated as if no such variation occurred. This holds especially of words that vary between open and close long *e*. Thus, Chaucer rimes *eke* 'also' now with *seke* 'seek', now with *breke* 'break,' a clear indication that in his dialect the strest vowel of *eke* might be either close or open. In other words, Chaucer's usage shows that in the 14th century, just as today, a given Englishman might habitually utter a given word now with one, now with another sequence of phonemes. Such a variation can hardly be called dialectal without straining the meaning of that term, even though the variant pronunciations may originally have belonged to different dialects. In old-fashioned terminology words that showed this variation between close and open long *e* in the strest syllable were said to have unstable *e* (the term is Skeat's). In the MED no attempt is made to mark this *e* as such. In the Prefatory Note we are told that the *ē* of an entry form may stand not only 'for the long open mid-front vowel' but also 'for a mid-front vowel that is open in some dialects but close in others'. Since in fact the variation occurs WITHIN the dialect (Southeast Midland of circa 1400 A.D.) upon which the dictionary is based, it will hardly do to dismiss the matter in this way.

In words like *ende* 'end' we have the complication that the strest vowel, which might be either long or short, was close when long but open when short. The editors mark the *e* here with both macron and breve, thus indicating the variation in quantity, but they ignore the variation in quality; indeed, the want of a point under the vowel sign may well lead the unsophisticated user of the dictionary to conclude that long *e* in this word was open in quality. The Prefatory Note ought to have included a statement covering cases of this kind.

But I do not wish to end this review without recording my full appreciation of the value of this dictionary for medieval English scholarship. We have here a vast array of quotations carefully gathered and carefully ordered under accurate definitions by a group of sound scholars, led by a man who not only 'knows his stuff' but also knows how to get things done. But for Kurath's leadership the great work in all likelihood would not have been finished within any foreseeable time. Thanks to him and his faithful collaborators, the long-awaited and long-needed dictionary of Middle English is now in sight. Indeed, its first fascicle lies before us. All medievalists who work in the English field, and all linguists

interested in this field, owe to Kurath, Kuhn, and the rest, and to the University of Michigan, a debt so great that it is beyond calculation. In the name of the many who will use the MED day by day the rest of their scholarly lives, let me congratulate the University, the Press, and the staff on a monumental accomplishment.

Morphologie du japonais moderne. By CHARLES HAGUENAUER. Vol. 1: Généralités, mots invariables; pp. vi, 425 (impression photomécanique). Paris: Librairie C. Klincksieck, 1951.

Reviewed by JOSEPH K. YAMAGIWA, *University of Michigan*

The author of this volume is professor of Japanese at the École des Langues Orientales, and director of studies at the École pratique des Hautes Études. The work is part of a larger plan. Vol. 2 will deal with the enclitics or particles, the 'mots variables' (adjectives and verbs), and the suffixes. Together, the two volumes will form the second part of the author's *Cours de langue japonaise moderne*, which will include also a volume entitled *Phonétique* and another entitled *Structure de la phrase*. These works are particularly welcome for the insight they provide into a European scholar's view of the Japanese language.

In his opening remarks, the author tells us (Avertissement, 1) that he has 'resolutely turned his back on all [previously written] grammars of Japanese, first as concerns the plan of his work and secondly as concerns the explanations proposed for particular forms and for the structure of the Japanese language'. He has 'occupied himself above all with the task of making the student feel that he is involved in a very special linguistic domain; he hopes that he has rid himself of grammatical notions which are of value only in his mother tongue, and has exerted himself to understand directly the linguistic facts that were placed before him, without recourse to these grammatical notions.' Further, he has 'tried to describe the Japanese language in conformance, as far as possible, with its own genius and its true nature. He has thus renounced the terminology which has remained in use in the grammars ... Deploring the fact that the majority of grammarians have disfigured the Japanese language by forcing it into an inappropriate framework, that of grammars which are valid for the Indo-European languages, he has attempted to give to the Japanese language an analytic description which is in better agreement with its "Altaic" character.'

After some further remarks on the principles underlying his work, the system of transcription used in his volume, and the kind of accentuation found in the Japanese language, the author devotes approximately 60 pages to a general discussion of the different categories and subcategories of Japanese words, and passes on to a more specific explanation of the different kinds of 'mots invariables'. These, he tells us, consist of nominal words, determinative words, honorific terms, number nouns, and numeral specifiers; further, uninflected words (and other words figuring as uninflected words) used in prepositional, adverbial, or conjunctive functions; onomatopoeic forms and other 'descriptive semantemes' (like *hin* in *himpin* 'frequency') borrowed from Chinese and used in reduplicated combinations; interrogative particles, the 'usual' interjections, the 'mots

injurieux' (pejorative and imprecatory), swear-words, and, finally, the 'indices for plurality'.

American linguists familiar with the structure of Japanese will shudder at the seeming hodgepodge of forms grouped under 'mots invariables'. This comes, however, from the fact that the author has used the word 'morphologie', as he says (2), 'in a very large sense': 'We understand by it even a *morpho-syntaxe*, that is to say, a study of words or signs as much from the point of view of their form and classification (forms of signs, categories of signs: *semantemes*, *morphemes*, and all other grammatical units that can be indicated) as from their functions and delimitations (behaviors establishing their interrelationships and the mechanism of these relationships; compatibilities and incompatibilities; point and manner of combining, etc.).' The 'mots invariables' (14) are 'all *semantemes* which can be cited and understood in their isolated state'. They suffer no changes in form, regardless of the suffixes that may be attached to them, and 'the native speaker conceives of them as full forms, not to be modified in usage, and, rightly or wrongly, as being simple and irreducible'. The *semantemes* are thus *morphemes* with 'full' meaning value, opposed to *morphemes* with grammatical or functional value. The latter for Haguénauer include *enclitics* (particles) and uninflected suffixes, which are only '*outils grammaticaux*', and which 'constitute the pieces, important to be sure, in the machinery of the language ... coming out of the shadows to become means of expression only when they are added to the *semantemes*'. The 'mots variables' also are *semantemes*, and the inflected suffixes attached to them are grammatical devices similar to the *enclitics* and the uninflected suffixes.

The *semantemes* and the non-*semantemes* which are thus set up are presented by Haguénauer in an interesting notation, in which all the morphs in a word or phrase are separated by means of hyphens and periods or otherwise. Thus the sentence meaning 'I've troubled you (with my visit; please excuse me)' appears as follows (147): *o-*ža.ma iłáš.i-máš.i-ta*. Here, to use Haguénauer's terms, the honorific term *o* is separated from the nominal word **ža.ma*; the asterisk indicates that *ža.ma* is a borrowing from the Chinese language, and the period between *ža* and *ma* merely separates the syllables. A period, again, separates the verb root *iłáš* from the thematic vowel *i*; the two together, forming a suspensive-formative base, are separated by a hyphen from the suspensive-conjunctive base (*-máš.i-*) of the auxiliary verb *-mas.ű*, which in turn is separated by a hyphen from the suffix *-ta*.

Enclitics (particles) act like suffixes, separated by hyphens from the forms they follow. Two *enclitics* or suffixes coming together are also separated by hyphens. A double asterisk indicates an archaic form, restored or attested. Whenever a word is made up of a native and a foreign element, the native part (provided it does not appear as the first element) is indicated by a small raised circle. A raised triangle indicates a form of the classic language or a survival from this language which is now felt to be obsolete. When a *semanteme* is made up of two or more elements which are themselves compounded, each of the elements is separated from the preceding by means of a small circle placed at the level of the line. A full slant or a raised slant indicates pauses and isolates 'logical

groups'. A prime indicates the loss of a vowel or consonant. The vowels *u* and *i* are marked with a breve when devocalized.

Further, every *r* has a dot over it, possibly to indicate that the Japanese *r* differs from the French. In transcriptions like *on.gak'.kwai* 'concert' and *wo* (accusative particle), Haguénauer's transcription seems to have partial reference to older pronunciations, and, in *on.gak'.kwai*, to the older kana orthography current up to the end of the war. He laments the fact that modern orthography confounds *zu* and *dzu*, *zi* and *dzi* (both pairs now homonymous), and asserts that **dzi* as the first syllable in the word for 'dictionary' is incorrect, the only proper pronunciation being **zi* (6). The word *hō* 'direction, side' is declared to be a monosyllable (13), though most linguists would call it dissyllabic (*ho-o*). In a passage discussing accentuation in Japanese (8-10), Haguénauer tells us that the word *tō* 'door' is pronounced on a mid pitch; the prevalent accentuation appears to be low. Except in this passage, Haguénauer makes no use of accent marks.

Despite its complexity, Haguénauer's system of transcription succeeds remarkably well in showing the etymological components of a word or phrase. The units that evolve, separated by hyphens, periods, circles, and spaces, serve to distinguish the semantemes and non-semantemes in which Haguénauer is interested. As he suggests (50), his transcription permits a quick distinction between consonant verbs (*kir.u* 'cut') and vowel verbs (*ki-ri.u* 'wear'). Some of the most interesting passages in his work are those in which he uses his transcription in speculating on the derivation of particular words and phrases.

American linguists will miss the customary distinction between bound forms and free forms that has been found so useful in descriptive studies on this side of the Atlantic. Haguénauer's concern with semantemes leads him to include among the 'mots invariables' the honorific prefixes *o-* and *go-*, the onomatopoeic and other 'descriptive semantemes', and such 'indices for plurality' as *-nado*, *-tači*, *-ŕa*, *-gata*, *-bara*, *-domo*, *-hai*, *-šū*, *-tō*, *moŕo-*, *šo-*, *sū-*, and *un.un*. Most of these forms have such meanings as 'and others, and so forth, various, several'. As we have seen, Haguénauer makes the confident statement that the native speaker conceives of these as 'full' forms. But if the feelings of a number of native speakers were actually to be consulted, one wonders how many of these forms would actually emerge as semantemes.

The author's subdivisions of 'mots invariables' thus suggest a thesaurus in which words of similar meaning are brought together. Most American linguists who have worked with Japanese would question the validity of an analysis in which five functions are found for the 'mots de qualité' (adjectives, 39-41), four bases for consonant verbs (47-8), three for vowel verbs (48-9), and four each for the 'hybrid' verbs *k.u-ri.u* 'come' and *s.u-ri.u* 'do' (51). But we must await the sequel to the present volume for Haguénauer's full description of these inflectional forms.

Haguénauer's work is in part directed at the French student working his way toward a practical knowledge of Japanese. The evidence is plentiful. Examples are the dot over the *r*, and the suggestion (7) that syllabic *n* is not to be pronounced with the vowel which precedes it, like a French nasalized vowel. French, says the author, has no honorific terms and no numeral specifiers such as are found in Japanese, and Japanese has no adjectives, relative pronouns, and articles

such as are found in French (14, 17). 'The necessities of pedagogy imply a minimum of rules to which the student may adhere and tie the things he knows' (28). The student reader's attention is directed to the quasi-personal pronouns (122-9) and to the 'mots invariables' and the forms figuring as 'mots invariables' which are used in prepositional, adverbial, or conjunctive function (22-4, 235-364). Speaking of the 'mots invariables' which take on prepositional functions, even Haguénauer admits (23) that 'it is only in translation that the *semantemes* in question figure as prepositions'. The difficulty with setting up a category like 'modes of expressing the comitative' (244-8) is that forms as diverse as *is.só* and *tomo*, which are nouns, and *tsur.e-te* and *sitaga'.e-te*, which are verb forms (gerunds), are classified in the same category.

Various other passages suggest the ways in which the French indefinites, the French conjunctions, and the French forms *en*, *y*, *on*, *se*, and *jamais* are translated into Japanese, or (106-7) how the Japanese determinatives may be rendered in French. Haguénauer strays even farther from a linguistic description of Japanese when he discusses the way in which loanwords are written (78-9), and farthest when he explains—very delightfully, to be sure—the use of the abacus (229-33). Many teachers of Chinese and Japanese will disagree with him when he states that the Japanese script should be learned together with the language, and that the student should never acquire a *semanteme* without learning the character which corresponds to it (3).

The kind of work that Haguénauer has produced may or may not illustrate a French view of structure. Surely those of us in this country who have written about the Japanese language must also have had our view of the subject-matter colored by our American-English background; and some of the features described by Haguénauer—for instance the adverbial use of nouns—are paralleled in many grammars written in English. A reading of his book suggests that it might be of considerable anthropological interest to compare several grammars of a given language written by linguists of different cultural backgrounds. Haguénauer's work is particularly interesting when viewed in this light; some parts of it, for instance those that have to do with the etymology of the *semantemes* (75-8), are even absorbing. The author's references to Chinese and Korean etymons, to Ryūkyūan cognates, and to dialect forms are consistently pertinent; his chapter on the numeral specifier is encyclopedic in its coverage; and his collection of 'mots invariables' in prepositional, adverbial, and conjunctive functions and of the onomatopoeic and descriptive *semantemes* is prodigious in scope. For these and other listings the reader is grateful; but the groupings are nearly always of semantically related forms, not of forms related by any linguistic criterion.

Unkhoswe waanyanja. By BENNETT E. MALEKEBU, ed. by GUY ATKINS. (Annotated African texts, No. 1: Maanja.) Pp. 124. Capetown and London: Oxford University Press, 1952.

Reviewed by DAVID L. OLMSTED, *Yale University*

This book is the first volume in a projected series to be issued by the School of Oriental and African Studies, University of London. Malekebu has written the texts, which deal with the system of guardianship among the Maanja

people of Nyasaland; Atkins has supplied the notes. There is little to be said about the texts by one who is ignorant of their language, except that the subject will be of considerable interest to the ethnologist. For that reason the texts achieve a double utility.

The most serious defect of the book is the lack of a glossary. Translations of some items are provided in the notes; but both students of the language and general readers will be considerably impeded in deciphering the text by the necessity of switching from Scott's or some other dictionary to the notes and back to the texts. Perhaps the editor expects his students to have at hand some more recently constructed Mapanja reading-aid; he does not tell us what it is.

The notes and preface do not further identify the 'Southern Nyanja' people, nor do they hint at a classification of their language or make detailed reference to any other work on the subject except the editor's *Suggestions for an amended spelling and word division of Nyanja* (International African Institute, Memorandum 25, 1950). Apparently the language in question is Johnston's language 61a Ci-mañanja of his (geographically classified) South Nyasaland group, in the Bantu subfamily of Greenberg's Niger-Congo family.

The assumptions underlying the orthography are veiled in obscurity, but occasionally a corner of the veil is lifted. The reader learns that the examples *fwáá* 'empty' and *fwá* 'full' 'show how necessary it is to take account of tone level and vowel length in exclamatory particles' (115). 'A straight line over a vowel is sometimes printed where a long vowel which occurs in speech is not written as a double vowel in the orthography' (65). Again, '*kalambulébwóló* is written as one word because it never takes concords of Class 5 [concords of Class 5 are not otherwise identified anywhere in the book]. If it did, *bwóló* would be hyphenated instead of joined' (124).

The grammatical notes are in the reviewer's opinion the greatest disappointment in the book. Terms from traditional Western grammar ('nominals of the agent' 66), picturesque analogies ('word building elements' 66), and cryptic abbreviations ('INV' 89, 'DNV' 91) are stirred in with references to some grammatical analysis, the existence of which can only be guessed at by the reader ('*wo* is a grammatical element of Class 3 [what Class 3?] and illustrates a semi-dependent usage' 70). The result is a potpourri of savory terms, without a hint of how they are to be identified with reference to the forms of the language, i.e. defined.

It is pleasant to be able to say that the notes not concerned with technical linguistics contain much that is of value. Aside from one or two dubious generalizations ('In a matrilineal society a daughter is a greater asset than a son' 81), there is creditable information on dialect differences (70), loanwords from Yao and English, and kinesic data associated with the language (68), and there are many valuable statements of correlated linguistic and non-linguistic cultural items. The kin terms should be a welcome addition to the information now available.

In brief, it is hoped that the series will continue; that the non-specialist will be given his bearings by fuller references, a glossary, and more careful linguistic statements; and that the valuable ethnolinguistic annotations will again accompany the texts. A final word about the printing and format: excellent.

Klinische und sprachwissenschaftliche Untersuchungen zum Agrammatismus.

By F. PANSE, G. KANDLER, and A. LEISCHNER. (Arbeit und Gesundheit: Sozialmedizinische Schriftenreihe aus dem Gebiete des Bundesministeriums für Arbeit, NF Heft 48.) Pp. 72. Stuttgart: Georg Thieme Verlag, 1952.

Reviewed by DAVID L. OLMSTED, *Yale University*

This monograph, the result of cooperative work by two psychiatrists (Panse and Leischner) and a linguist (Kandler), is a welcome addition to the literature on aphasia. If the authors have not succeeded completely in integrating linguistic theory and method with the methods traditionally used in the study of this disorder, the reason may be sought in the relative novelty of the goal and in the very great difficulties presented by the field itself.

As long ago as 1913, Pick¹ pointed out the desirable consequences of cooperation between psychologists, neurologists, and linguists in the study of aphasia; later investigators have usually paid lip-service to the notion. Jakobson's monograph,² which twelve years ago surveyed the literature from a linguistic point of view and advanced some interesting hypotheses relating to aphasia, seems not to have received the attention it deserves. The neurologist, who is actually on the firing line in the study of aphasia, is busy attempting to alleviate human misery, and cannot be expected to master all the disciplines that might conceivably be relevant to what is only one of his problems. By bringing a linguist into his group of investigators, Panse has therefore given linguistics a chance to prove its own relevance to aphasia studies.

The monograph is noteworthy in adding to the literature a relatively full report of a single case, a desideratum in a field so far characterized by inchoate theory and method. Besides linguistic protocols, the book includes the results of various medical tests and observations, which the reviewer is not competent to evaluate.

The authors recognize the necessity of operationally defining or redefining certain concepts traditional in studies of aphasia. They do not always succeed, but the attempts strike the reviewer as laudable and not without value. For example, the disturbances that constitute agrammatism are defined as 'malperformances in the grammatical province of language'. The grammatical province is 'die Schicht der Wortfügungen (Bildung von Wörtern, Wortgruppen und Sätzen, dazu die "Fügungsseite" des einzelnen Wortes) mit ihren inhaltlichen Werten, die äusserlich greifbar werden an Wortstellung, phonetischer Gliederung, Partikeln und Flexionsformen.' There may still be questions concerning the definition of some of the terms used to explicate 'grammatical', but these seem of less importance than the question whether a linguist—a specialist in the language concerned—could RECOGNIZE such malperformances upon hearing them. Since the reviewer is inclined to believe that the answer is Yes, he views the quoted formulation as a practical advance, although perhaps leaving room for final improvements.

As Panse and his colleagues recognize, it follows from their definition of agrammatism that the disorder is not, as it was once thought to be, peculiar to 'in-

¹ A. Pick, *Die agrammatischen Sprachstörungen* 182 (Berlin, 1913).

² Roman Jakobson, *Kindersprache, Aphasie und allgemeine Lautgesetze* (Uppsala, 1941).

flecting languages', and that the specific behavior said to constitute it must be redefined for each language. Though they subscribe to some unsubstantiated notions about Chinese, perhaps derived from the sources they consulted,³ the authors rather neatly illustrate the linguistic relativism of the syndrome with comparative material from Chinese and Abkhasian. Here, as elsewhere, they might have saved themselves some work by a closer reading of Jakobson's monograph.

The chief material taken as the corpus for linguistic analysis is the patient's spontaneous speech, a method that has considerable disadvantages, since there are many more variables uncontrolled than in the standard test situations. Still, the monograph makes an important point: for this patient, no grammatical 'possibility' is fundamentally or completely 'blocked.' 'It is thus not necessary to postulate an independent psycho-physiological grammatical or speech-combination model which was disturbed as such.' Agrammatism is a matter of probabilities, not an all-or-none proposition. This conclusion, if substantiated elsewhere, should cast doubt on the inclusiveness of the standard tests—quite apart from the question whether they actually test what they purport to test.

Spontaneous speech (to paraphrase the monograph) may show less mal-performance than speech in a test situation, because of the greater number of cues from the contextual redundancies of the language. The authors emphasize the desirability of continuing to study the spontaneous speech of aphasics, and call for reports on aphasic speakers of non-Indo-European languages. They record their views on localization, a perennial problem in this field; but here again the reviewer must disqualify himself.

There may be some differences of opinion with respect to the interpretation of the behavioral data. The trail of aphasia studies is cluttered with discarded intervening variables and whitened with the bones of hypothetical constructs;⁴ both are illustrated in this monograph. 'Autonomous weaknesses of the acoustic-gnostic function' are found unuseful, as are Goldstein's Gestaltist formulations. In their place we find such hypothetical constructs as 'defective mnestic dispositions' and 'die mangelhafte Weckbarkeit der grammatischen Gesetzlichkeit des Sprachmaterials'.

In sum, the monograph is to be welcomed. The authors have contributed to the store of data, to the theory of aphasia, and to the observational techniques suggested by the theory. It is to be hoped that their work will point the way to wider utilization of linguistics in such studies.

³ G. von der Gabelentz, *Anfangsgründe der chinesischen Grammatik* (Leipzig, 1883); E. Schmitt and Y. Lou, *Einführung in das moderne Hochchinesisch* (Shanghai, 1939).

⁴ Cf. K. MacCorquodale and P. E. Meehl, On a distinction between hypothetical constructs and intervening variables, *Psychological review* 55.95-107 (1948).

NOTES

THE AMERICAN COUNCIL OF LEARNED SOCIETIES, through its Committee on the Language Program, is directly or indirectly responsible for encouraging, planning, and guiding many of the projects in which American linguists are currently engaged. The members of this Committee are appointed by the Board of Directors of the ACLS; at present they are Albert H. Marckwardt (chn.), Norman A. McQuown (secy.), Bernard Bloch, John B. Carroll, Stephen Freeman, Archibald A. Hill, Martin Joos, and Henry Lee Smith Jr., with Mortimer Graves as ACLS liaison officer and John Kepke as business manager. The following statement sets forth the Committee's view of the situation in which it works and of its own specific duties.

LANGUAGE STUDY AND AMERICAN EDUCATION

1. **The need for competence in languages.** The past fifty years have witnessed the sudden transformation of the United States from an insular and continental to a world power. To quote from an earlier publication of this organization,¹

The product of American industry spreads all over the world. Wherever there is a paved road there is an American automobile; American oil is produced wherever there is oil and used wherever oil is used. American banks have branches and connections in every significant foreign city. Hardly a remote corner of the globe but knows the American missionary, the American school, American generosity. ... Half the world's trains run on American rails. No region is too remote to be the concern of American diplomacy. And all too frequently American armed forces must ply their trade in lands and among peoples whose very names would have been unknown to an earlier generation.

One would suppose accordingly that many Americans would be equipped with scientific and detailed understanding of these multifarious cultures, that the United States would lead the world in the study of foreign lands no matter how distant, that no society could be named for which there was not an American expert, and that the American academic structure would reflect this world perspective. Unfortunately, a true picture is almost the reverse of this. ... Study of many culturally and strategically important peoples finds no place whatever in American universities and colleges.

The importance of language study in meeting this situation is clear. Consider, for example, that Asian thought alone is published in about twenty different languages. To understand Asia, the United States needs facilities for teaching and learning these languages. This in turn requires tools—textbooks, graded readings, recordings, dictionaries—which at present are available for scarcely a single Asian language other than Japanese and Chinese. In linguistics as in the natural sciences, the ultimate production of tools depends on basic research: in this case, the scientific analysis of the languages in question. Until we have reliable and reasonably exhaustive descriptions of Asian languages, we cannot provide the tools that Americans must use in taking the first step toward an understanding of Asian cultures. And until that step is taken, the United States can make no real progress in meeting its obligations to the world.²

If it is important to build a linguistic bridge to the diverse cultures of Asia, it is no less important to maintain and improve our understanding of Western Europe. We share a common cultural tradition with the peoples of that area; it is chiefly among these peoples that

¹ Mortimer Graves, *A neglected facet of the national security problem 1* (Washington, 1950, reprinted 1952).

² Paraphrased from Graves, *A program for the improvement of American understanding of Asian civilizations*: Memorandum No. 2, Basic implementation for language and area studies 8-9 (Washington, 1951).

we can hope to find a belief in democracy similar to our own; and it is only with their sympathetic cooperation that we can help to maintain the free world.

Though we sometimes speak of the American academic structure as centered upon Western Europe, we can no longer cheerfully assume that the student who emerges from it has acquired an effective command of any continental West European language, a working knowledge of any West European literature, any systematic information about West European political, economic, or social conditions, or indeed any appreciation of the differences between various West European attitudes and his own. The average student is naively unaware that people of other backgrounds do not think like Americans, do not react like Americans, and have a sense of values often quite different from the American. Add to this, that with the decline of Greek and Latin studies, students have lost the key to an understanding of that classical civilization from which their own has developed. American learning, like many other aspects of American life, is impoverished by its orientation in a single language and a single culture.

Everyone knows that our society is today many times more complex, many times more intimately bound up with other nations than it was two centuries ago. Behind the purchase of a spool of thread, a pair of shoes, or a newspaper lies the planned labor of hundreds of men throughout the world. The structure of government, from the town to the nation, is complicated beyond the dreams of our grandfathers. And nowhere is the intricacy of present-day economic structures more obvious than in our own country, where more than a hundred and fifty million people are involved in a daily network of communication and co-operation.

At the same time we are traditionally committed to the faith that this complex social structure must operate by democratic means—that every member of the society has a share in its control. The strains and tensions that inevitably arise from the interaction of so many people can be resolved only through communication at its most sensitive—that is, through language. Our democracy can survive only with a literate and articulate body of citizens, who can react intelligently to essential communications and can adequately express their wants, their attitudes, and their opinions. Such a command of language implies the ability not only to speak and write effectively but to read and listen critically. With ordinary correctness in the schoolbook sense it has little or nothing to do; but it requires a practical knowledge of the standard language, and of its regional variations, far beyond anything that any nation has yet attained.

Nor can we overlook the enormous changes that have recently taken place in the vehicles of communication. The radio, the movies, television, and the use of public-address systems make it possible to disseminate voices and gestures—to say nothing of the printed word—on a scale that would have seemed fantastic only a generation ago.

The development of these mass media has intensified our educational responsibilities. Properly employed, such media are instruments of great social utility: they can alleviate suffering, enlighten ignorance, and heighten the significant experience of every individual. Wrongly used or manipulated by the enemies of democracy, they are a menace to our social structure. But they cannot become a menace so long as their audience is critical, informed, and questioning. To make it so is the responsibility of American education. Students must be trained to search for the attitudes and presuppositions implied in the messages that come to them through these mass media: to weigh them dispassionately, to recognize every semantically slanted term, to test the aptness of every analogy.

The task is not easy. It will not be accomplished by new courses in movie appreciation and radio listening; certainly it will not be furthered by relaxing our standards of attainment in the high school and the college. Nor will a lighthearted or a despairing acceptance of popular anti-intellectualism help the teacher to discharge his responsibility to the country. We can avoid the latent danger, but only by insisting firmly and repeatedly on scientific rigor in place of easy popularization, and by emphasizing always an enlightened view of the nature and functions of language. This is a challenge that we dare not fail to meet.

Language is in every known society the communicative process par excellence. The

history of civilization is the history of a steady increase in the radius of communication. If we are to help students to understand human behavior—their own as Americans and that of at least one other society—we must begin with the study of language; for it is through this system that all the other systems of culture are most adequately observed and transmitted. To clarify the function of language in any social group is to make a start in teaching cultural relativity and hence in dispelling harmful ethnocentric attitudes. Linguistics and language learning are central in the integration of the social sciences and the humanities; and at the very core of every student's total language experience is the command of his native language.

2. The attainment of competence in languages. To raise the level of our nation's linguistic competence requires more than good intentions; it requires a method. We must (a) extend our understanding of linguistic principles through basic research; (b) apply the resulting concepts to particular areas; and (c) cooperate with those who can put life into such applications.

(a) **THE BASIC SCIENCE.** The scientific study of language has developed recently and rapidly; in the United States especially it has been carried to a high level of competence. Basic research in this field involves the descriptive analysis of language in its various aspects, and of particular languages, past and present, for their own sake and in relation to the nonlinguistic culture of the people who speak them. Much has been accomplished, but vastly more remains to be done; even the basic principles of linguistics need clarification. Only a small fraction of the languages of the world have been adequately described; the rest, still to be dealt with, include many that are on the point of extinction. To record these languages before they die, we need a significant increase in the number of trained field workers. And for this in turn we must extend the teaching of linguistics both to graduate students and to undergraduates; we must disseminate the products of research more effectively; and we must give increased opportunities for study and fruitful work to those already competent in the field.

(b) **PRACTICAL APPLICATION.** Learning to work with a language is not essentially different from learning to control any other intricate structure. The first essential is to know how it operates: to have access to an accurate and comprehensive description. But as we have already observed, most of the languages of the world are still unfamiliar to us, and the familiar languages of Western Europe are too often taught in terms of a grammatical apparatus designed for Greek and Latin. The traditional eight parts of speech, recognized in some but not all Latin grammars, are a poor fit for the categories of English or French; in many languages, even such a basic Latin dichotomy as that between adjective and verb is irrelevant. And the confusion in most Latin grammars between tenses (past, present, future) and aspect (perfect, imperfect), implied in the conjugation of Latin verbs, is not helpful in making clear the difference in English between *he went* and *he has gone*.

To teach a foreign language effectively it is necessary to compare, point by point, the structure of the language to be learned and that of the learner's native tongue. Only in this way is it possible to determine what is difficult and what is easy, and so to devise an efficient order of presentation. A teacher need spend no time training Germans to pronounce French *u*, whereas English-speaking students need careful drilling—because, obviously, the sound occurs in both French and German but not in English. The two-gender system of Spanish nouns is fairly easy for a Frenchman, who has a similar system in his own language; but the speaker of English finds it an irrational nuisance.

Successful instruction in a foreign language requires textbooks, graded reading materials, and students' dictionaries, each constructed with the linguistic orientation of the American learner in mind. It requires the teacher, or at least the writer of the textbooks, to restate the results of the structural analysis of English and the foreign language in terms that will be comprehensible to every student. If the student is to extend his command of any language, including his own, and especially if he is to integrate his knowledge of the language with what he learns about the society that speaks it, the teaching materials must be assembled with tact and care, and presented in a manner at once ingenious and urbane, analytical and esthetically sensitive. The materials used in the teaching of languages not

only will have to be re-created almost completely within the next twenty-five years, but will have to be designed on patterns quite different from those to which we have grown accustomed.

(c) **IMPLEMENTATION.** The last half-century has seen a steady decline in the prestige of languages as a school subject, in the importance of their position in the curriculum, and in the skill of students who have been subjected to language instruction. Among the many reasons for this, the most important is that languages have not been well taught. Students come out of classes in the foreign languages unable to read, to converse, or to compose. And even the student of English, though he may master a special dialect in which to write his themes, forgets or discards it as soon as he leaves the classroom. Languages thus blunderingly presented have naturally failed to provide a key to the understanding of our own or any other culture.

As a result, the study of language is now viewed with hostility by the various professions, the public at large, and the administrators who have in their charge the construction and operation of school curricula. The task of reestablishing confidence in language study is difficult and complex, but there is one simple way to begin it: to demonstrate how languages can be effectively taught at various levels of education and how they can serve as keys to our own and other civilizations. If we succeed in showing this, the battle will be more than half won; but we cannot show it without the basic research and the techniques of application already mentioned.

3. Language study and the ACLS. For many years the American Council of Learned Societies has recognized an increase in the range and effectiveness of language instruction as a primary educational need. Shortly after the United States entered the war, the sudden demands for competence in many languages hitherto scarcely heard of led the Council to establish the Intensive Language Program. The committee in charge was at first concerned to establish intensive courses in colleges and universities throughout the country, but almost immediately began to develop teaching materials for familiar languages as well as for those less commonly taught. At the end of the war the need for trained linguists grew less pressing, and funds for the development of personnel were correspondingly reduced. Nevertheless, the ACLS COMMITTEE ON THE LANGUAGE PROGRAM has continued to further both basic research in linguistics and the application of its results to the teaching of languages. As a consequence of recently renewed steps toward preparedness in America, and of the development of language-teaching activities in American cultural centers abroad, the COMMITTEE has again enlarged the scope of its operations.

In the light of our situation as described at the beginning of this report, the COMMITTEE feels that an explicit formulation of its aims and activities may be of general interest. Such a formulation is here attempted.

FUNCTIONS OF THE ACLS COMMITTEE ON THE LANGUAGE PROGRAM

1. To encourage research in all aspects of linguistic science.
2. To provide aids to the publication of sound linguistic research.
3. To develop and encourage instruction in linguistics for both graduate students and undergraduates.
4. To develop opportunities for integrating linguistics with the study of all other fields of human behavior.
5. To apply linguistics to the teaching of foreign languages at all levels of instruction.
6. To apply linguistics to the teaching of the native language from literacy to literature, and to the teaching of all other relevant phases of culture.
7. To develop training programs to acquaint teachers with the principles and possibilities of linguistics.
8. To develop in the public an awareness of current activities among linguistic scientists.
9. To consider and present opportunities for the employment of personnel trained in linguistics.
10. To bring to the attention of administrators the importance of linguistics for present-day education.

EDITH FRANCES CLAFLIN, a Foundation Member of the Linguistic Society of America, died at her home in New York on Thursday, 5 March 1953, aged 77 years. To her family, her many friends, and her colleagues in the domain of classical linguistics, the loss is a heavy one.

Edith Claflin was born on 6 October 1875 in Quincy, Mass., the third child and second daughter of Frederick Allan and Adelaide Avery Claflin.³ After attending the Coddington public school of Quincy and the Thayer Academy of Braintree, she entered Radcliffe College, where she earned her B.A. in three years, graduating in 1897 magna cum laude and Phi Beta Kappa. From the time that she began the study of Greek at Thayer Academy, she had made up her mind to become a philologist. With this object in view, she set herself to master the language so perfectly that—as she afterwards told her sister—she could not remember having ever made a mistake in it. Greek and Latin were her major subjects at Radcliffe, and later during two years of graduate study at Bryn Mawr College (1897–99). From 1899 to 1900, as Garrett European Fellow from Bryn Mawr, she worked at the American School of Classical Studies in Athens. During a visit to Boeotia she laid the foundation for her doctoral dissertation, *The syntax of the Boeotian dialect inscriptions*.

Unable to secure a university appointment upon her return from Athens in 1900, she became a teacher in various preparatory schools. Thus impeded in her research, she did not obtain her doctorate from Bryn Mawr until 1904. From 1907 to 1913 she taught at Monticello College in Alton, Ill.; from 1914 to 1916, at the Laurel School in Cleveland; and from 1916 to 1933, at Rosemary Hall in Greenwich, Conn., as head of the Department of Greek. For one semester, in 1926, she was Acting Professor of Latin at Indiana University.

The years of teaching in secondary schools laid a heavy burden upon her; but they gave her a fund a pedagogical skills that were to make her college teaching an easy task. In 1936 she was appointed Lecturer in Greek and Latin at Barnard College; and in the same year, in the Columbia University School of General Studies (then called University Extension), she introduced a course in Medieval Latin which continued to attract students to her even after she had retired from her position at Barnard in 1945—which continued to attract them, indeed, up to the time of her death. The enthusiasm of her more mature students was as warm and lasting as that which she had formerly aroused in younger learners.

Among scholars Edith Claflin was known as an authority on the Indo-European middle voice, particularly on the *-r*-forms in Latin and other languages. She was primarily a grammarian, competent in phonology and semantics but going into these aspects of language no more deeply than her subject required. However rigorous her methods, she always retained a vivid sense of the human situations in which language is used. Besides her devotion to the languages of classical civilization, she had also a strong love of English literature; one of her

³ Biographical details from the article in *Directory of American scholars* (Lancaster, Pa., 1951); from the obituary notice in the *New York Times* of Sunday, 8 March 1953; and from a letter kindly written by Edith Claflin's sister, Miss Charlotte Isabel Claflin of Buffalo. Several passages in my account are quoted almost verbatim from this letter.

less recondite articles is entitled *Latinisms in Shakespeare's diction*. Her principal scholarly publications are the following:

The syntax of the Boeotian dialect inscriptions (dissertation); Bryn Mawr College monographs, 1906.

The nature of the Latin passive, *AJP* 48.157-75 (1927).

The hypothesis of the Italo-Celtic impersonal passive in *-r*, *Lg.* 5.232-50 (1929).

Venetic *tolar*, Old Irish *canar*, and the Indo-European injunctive, *Lg.* 12.23-34 (1936).

The Indo-European middle ending *-r*, *Lg.* 14.1-9 (1938).

The voice of the Indo-European perfect, *Lg.* 15.155-9 (1939).

The text of Plautus, *Curculio* 260-261; *Epidicus* 61-62, *PAPA* 71.xxxi-ii (1940).

The middle verb *vidēri*, *Lg.* 18.26-32 (1942).

Videor as a deponent in Plautus, *AJP* 64.71-9 (1943).

Review of Dag Norberg, *Beiträge zur spätlateinischen Syntax*, *AJP* 70.211-7 (1949).

She enjoyed the company of her colleagues, and could seldom resist the lure of a professional meeting, annual or monthly. To keep in touch with current developments, she often made the trip to New Haven for meetings of the Yale Linguistic Club; and at meetings of the Linguistic Society she was for many years a familiar and prominent figure. (It is worth mentioning that in 1943 and '44 she was a member of our Executive Committee.) Besides the Linguistic Society she belonged to the Philological Association, the Mediaeval Academy, the Classical League, and the Classical Associations of the Atlantic States and of New England.

BERNARD BLOCH, *Yale University*

WE RECORD WITH SORROW the death of three other members of the Linguistic Society. Richard Jente, a member since 1941, died on 22 August 1952, aged 64; George Raymond Shipman, a member since 1942, died on 22 January 1953, aged 43; and Charles Grimm, a member since 1926, died on 4 April 1953, aged 58.

THE WASHINGTON LINGUISTIC CLUB (see *Lg.* 23.186) has recently become active again after a lapse of several years. Officers for the year just ending were A. A. Hill, chairman; F. A. Rice, vice-chairman; and P. L. Garvin, secretary and treasurer. The Club holds regular monthly meetings during the academic year for the reading of papers. There are now more than fifty members; the average attendance at the meetings is about 40. Dues are \$1.50 per calendar year. All interested persons in Washington and near-by cities are invited to join the Club and attend its meetings; for information they should address P. L. Garvin, 1719 Massachusetts Ave., N.W., Washington 6, D. C.

EDGAR G. C. POLOMÉ calls attention to a misstatement in his article On the source of Hittite *h*. In *Lg.* 28.444 (1952), the last sentence of footnote 4 should

read as follows: 'Giuliano Bonfante quotes this view, without subscribing to it, in CP 39.52 fn. 4 (1944).'

THE STYLE RULES OF LANGUAGE have long provided that the titles of books and journals should be cited in roman type, with only the first word and the proper names capitalized. On the whole this practice has worked very well, and has contributed to the clean look of our pages; but sometimes it has proved inconvenient, as when a title occurs in a context where it can be misread as an ordinary part of the sentence. (Thus the words 'Sanskrit grammar' constitute a book title when we speak of Whitney's, but not when we speak of Wackernagel's.) To avoid such ambiguities and minor unclarities, the Editor has decided to experiment this year with a different convention. In the current volume of LANGUAGE, titles of books and journals (but not of articles) are cited *in italics*, though still without superfluous capitals. If the new practice proves more efficient than the old, it will be permanently adopted. (But contributors to the journal are asked to omit underlines as usual, since it is easier to add than to erase them.) Meanwhile, the Editor asks the indulgence of readers for inconsistencies observed during the period of transition. Old habits die hard.

PUBLICATIONS RECEIVED

This listing acknowledges the receipt of recent works that appear to bear on the scientific study of language. No book can be returned to the publisher, nor can the Editor promise that every book received will be reviewed in the journal. Reviews are published as circumstances permit, and copies are sent to the publishers of the works reviewed.

- Acta linguistica* 7.1-92 (1952).
Acta linguistica Academiae Scientiarum Hungaricae 1.1-482 (1951-52); 2.1-145 (1952).
 Budapest: Magyar Tudományos Akadémia.
Aegyptus 32.1-240 (1952).
Afrika und Übersee 37.1-48 (1952).
Anthropology today: An encyclopedic inventory. By A. L. Kroeber and others. (International symposium on anthropology, Wenner-Gren Foundation for Anthropological Research: The inventory papers.) Pp. xv, 966. Chicago: University of Chicago Press, 1953.
El Antijovio. By Gonzalo Jimenez de Quesada, ed. by Rafael Torres Quintero, with 'estudio preliminar' by Manuel Ballesteros Gaibrois. (Publicaciones del Instituto Caro y Cuervo, No. 10.) Pp. clxxiv, 637, with 11 plates. Bogotá, 1952.
An appraisal of anthropology today. Ed. by Sol Tax, Loren C. Eiseley, Irving Rouse, and Carl F. Voegelin. (International symposium on anthropology, Wenner-Gren Foundation for Anthropological Research.) Pp. xiv, 395. Chicago: University of Chicago Press, 1953.
Arabisches Wörterbuch für die Schriftsprache der Gegenwart. Ed. by Hans Wehr. Vols. 1-2, pp. xi, 986. Leipzig: Otto Harrassowitz, 1952.
Archiv für Orientforschung 16.1-196, i-viii (1952).
Archiv orientální 19.321-644 (1951).
Asia major NS 3.1-120 (1952).
L'aspect et le temps dans le verbe néo-grec. By Hansjakob Seiler. (Collection de l'Institut d'Études Byzantines et Néo-Helléniques de l'Université de Paris, Fascicule 14.) Pp. 171. Paris: Société d'Édition 'Les Belles Lettres', 1952.
Der Begriff 'rhēma' im Biblisch-Griechischen: Eine traditionsgeschichtliche und semologische Untersuchung; I. 'Rhēma' in der Septuaginta. By Eero Repo. (Annales Academiae Scientiarum Fennicae, Ser. B, Vol. 75.2.) Pp. 205 Helsinki, 1951.
Berceo 7.375-728 (1952).
Bulletin analytique: Philosophie 6.591-792, i-vi, 3-47 (1952).
Bulletin of the School of Oriental and African Studies 14.413-681 = Studies presented to Vladimir Minorsky (1952); 15.1-202 (1953).
Boletim de filologia 13.1-183 (1952).
Cahiers Sextil Pușcariu 1.1-453 (1952). Linguistique, philologie, littérature roumaines. Roma: Editions Dacia; Valle Hermoso: Cartea Pribegiei.
The classical weekly 46.49-64 (1952), 65-160 (1953).
A critical bibliography of the new stylistics applied to the Romance literatures 1900-1952. By Helmut A. Hatzfeld. (The University of North Carolina studies in comparative literature, No. 5.) Pp. xxii, 302. Chapel Hill, [N.C.], 1953 (1952).
La dialectología griega como fuente para el estudio de la migraciones indoeuropeas en Grecia. By F. Rodríguez Adrados. (Acta salmanticensia: Filosofía y letras, Vol. 5, No. 3.) Pp. 73. Universidad de Salamanca, 1952.
Dictionnaire française-lomongo (Lonkundo). By G. Hulstaert, M.S.C. (Annales du Musée Royal du Congo Belge; Sciences de l'Homme: Linguistique, Vol. 2.) Pp. xxxii, 466. Anvers: Éditions De Sikkel, 1952.
Diogenes 1.1-147 (1952). An international review of philosophy and humanistic studies: Quarterly publication of the International Council for Philosophy and Humanistic Studies. Anglo-American edition published by Intercultural Publications Inc., New York.
Études germaniques 7.241-320 (1952); 8.1-96 (1953).

- A grammar of the Macedonian literary language. By Horace G. Lunt. Pp. xv, 287, with folding map. Skopje, [Macedonia], 1952.
- Growth and system of the language of Dickens: An introduction to a Dickens lexicon, revised edition. By Tadao Yamamoto. Pp. [ii], 508, 8, 70. [Osaka, Japan]: Kansai University Press, 1952.
- Grundriss der akkadischen Grammatik. By Wolfram von Soden. (Analecta orientalia: Commentationes scientificae de rebus orientis antiqui, Vol. 33.) Pp. xxvii, 274, 51*. Roma: Pontificium Institutum Biblicum, 1952.
- Gwary kujawskie [Les dialectes coufiaviens]. By Zenon Sobierajski. (Poznańskie Towarzystwo Przyjaciół Nauk, Wydział filologiczno-filozoficzny, Prace komisji filologicznej, Tom 14, Zeszyt 2.) Pp. iv, 128, with folding map. Poznań: Nakładem Poznańskiego Towarzystwa Przyjaciół Nauk, 1952.
- An historical dictionary of German figurative usage. By Keith Spalding. Fasc. 1-2, *A to aufheben*, pp. i-vii, 1-88. Oxford: Basil Blackwell, 1952.
- Indogermanisches etymologisches Wörterbuch. By Julius Pokorny. Fascicle 7, *ker- to leizd-*, pp. 557-672. Bern: A. Francke AG Verlag, 1953.
- Institut Royal Colonial Belge 23.1-276 (1952).
- International journal of American linguistics 19.1-84 (1953).
- The interpreter's handbook: How to become a conference interpreter. By Jean Herbert. (Université de Genève, Faculté des Lettres, École d'Interprètes.) Pp. vi, 113. Genève: Librairie de l'Université (Georg & Cie. S.A.), 1952.
- Intonation of conversational English (educated southern British). By Wiktor Jassem. (Travaux de la Société des Sciences et des Lettres de Wrocław, Seria A, Nr. 45.) Pp. 121. Wrocław: Nakładem Wrocławskiego Towarzystwa Naukowego, 1952.
- An introduction to a survey of Scottish dialects. By Angus McIntosh. (University of Edinburgh: Linguistic survey of Scotland, Monographs, No. 1.) Pp. xii, 122. Edinburgh etc.: Thomas Nelson and Sons Ltd. (for the University of Edinburgh), 1952 (1953).
- Isländisches etymologisches Wörterbuch. By Alexander Jóhannesson. Fascicles 2-3, pp. 161-480, *gel- to demā-*. Bern: A. Francke AG Verlag, 1952.
- Italica 29.215-78 (1952); 30.1-64 (1953).
- Izvestija Akademii Nauk SSSR 11.393-576 (1952).
- Jahrbuch für kleinasiatische Forschung 2.225-352 (1953).
- Jezik 1.1-96 (1952). Časopis za kulturu hrvatskoga književnog jezika; Zagreb (Yugoslavia).
- Journal de la Société des Américanistes NS 41.1-220 (1952).
- Journal of the Oriental Institute 2.1-104, 33-40 (1952). M.S. University of Baroda, Baroda.
- The journal of the Polynesian Society 61.1-175 (1952).
- Klinische und sprachwissenschaftliche Untersuchungen zum Agrammatismus. By Fr. Panse, G. Kandler, and A. Leischner. (Arbeit und Gesundheit: Sozialmedizinische Schriftenreihe aus dem Gebiete des Bundesministeriums für Arbeit; NF Heft 48.) Pp. 72. Stuttgart: Georg Thieme Verlag, 1952.
- Kost' Michal'čuk. By Jurij Šerech. (Ukrains'ka Vil'na Akademijska Nauk, Serija: Ukrains'ki Včeni, No. 3.) Pp. 40. Winnipeg: Nakladom Ukrains'koj Vil'noj Akademii Nauk, 1952.
- Kultur und Sprache: Ein Festband dem 4. Internationalen Kongress für Anthropologie und Ethnologie gewidmet vom Institut für Völkerkunde der Universität Wien. Ed. by Wilhelm Koppers, with the collaboration of Robert Heine-Geldern and Josef Haekel. (Wiener Beiträge zur Kulturgeschichte und Linguistik, Vol. 9.) Pp. 511. Wien: Verlag Herold, 1952.
- Languages of West Africa. By Dietrich Westermann and N. A. Bryan. (Handbook of African languages, Part 2.) Pp. 215, with folding map in pocket. London, New York, Toronto: Oxford University Press (for the International African Institute), 1952.
- Die Lehnwörter des Sajansamojedischen. By Aulis J. Joki. (Mémoires de la Société finno-ougrienne, No. 103.) Pp. 395. Helsinki: Suomalais-Ugrilainen Seura, 1952.
- Leuvense bijdragen 42.1-91 Bijblad 1-53, (1952).
- Levende talen, Nos. 167, 168 (1952).
- Lexical number in Spanish nouns with reference to their English equivalents. By James E.

- Iannucci. (University of Pennsylvania, Romance languages and literatures: Extra series, No. 12.) Pp. xii, 80. Philadelphia: University of Pennsylvania, 1952.
- Lexis 3.1-164 (1952).
- Lingua 3.121-241 (1952).
- Lingua nostra 13.97-124 (1952).
- Le lingue dell'Italia antica oltre il latino. By Vittore Pisani. (Manuale storico della lingua latina, Vol. 4.) Pp. xvii, 354, with folding chart and two maps in pocket. Torino: Rosenberg & Sellier, 1953.
- Linguistica I (1941-1945). Ed. by John Lotz. (Acta Instituti Hungarici Universitatis Holmiensis: Series B, Linguistica, No. 1.) Pp. 56. Stockholm, 1952.
- Lud 39.1-806 (1948-51). Organe de la Société Polonaise d'Ethnologie; Kraków-Poznań, 1952.
- Le mètre phonétique III.30.19-49 = No. 98 (1952).
- Man 52.145-92 (1952).
- Manuel de langue roumaine: Grammaire, textes d'étude commentés avec index grammatical et glossaire. By Alain Guilleumou. (Les langues d'Europe orientale, Vol. 5.) Pp. 285. Paris: Librairie C. Klincksieck, 1953.
- Meijerbergs arkiv för svensk ordforskning 8.1-128, 9.1-151 (1952).
- Middle English dictionary. By Hans Kurath, editor, and Sherman M. Kuhn, associate editor. Part E.1, *ē* to *ēndelōnges*, pp. ii, 120. Ann Arbor: University of Michigan Press, 1952.
- A Minoan Linear B index. By Emmett L. Bennett Jr. Pp. xxiv, 119, offset. New Haven: Yale University Press (published for the Department of Classics), 1953.
- A modern Czech grammar. By William E. Harkins, assisted by Marie Hnyková. (Columbia Slavic studies; A series of the Department of Slavic Languages, Columbia University.) Pp. xi, 338, offset. New York: King's Crown Press, Columbia University, 1953.
- Modern Hebrew. By Eliezer Rieger. Pp. 156. New York: Philosophical Library, 1953.
- Der Name der Buche. By Wilhelm Wissmann. (Deutsche Akademie der Wissenschaften zu Berlin: Vorträge und Schriften, Heft 50.) Pp. 38. Berlin: Akademie-Verlag, 1952.
- Neophilologus 36.193-256 (1952); 37.1-64 (1953). Dertigjarig register, by W. J. Schuijt; pp. 108; Groningen, Djakarta: J. B. Wolters, [1953].
- Les neuf joies nostre Dame; A poem attributed to Rutebeuf. Ed. by Tauno F. Mustanoja. (Annales Academiae Scientiarum Fennicae, Ser. B, Vol. 73.4.) Pp. 90. Helsinki, 1952.
- Notes de grammaire rundi. By A. E. Neeussen. Pp. 24, mimeographed. Tervuren [Belgium]: Musée Royal du Congo Belge, 1952.
- Nueva revista de filología hispánica 6.109-208 (1952).
- Onoma 2.1-118, 1*-69* (1951). Bulletin d'information et de bibliographie; Comité International des Sciences Onomastiques, Centre International d'Onomastique, Louvain.
- Orbis 1.335-616 (1952).
- Orientalia 22.1-128, plates i-xxviii (1953).
- Orientalia suecana 1.95-114, plates i-xvi (1952).
- Participium universale im Slavischen. By Jurij Šerech. (Slavistica: Abhandlungen des Instituts für slavistische Forschung der Ukrainischen Freien Akademie der Wissenschaften, No. 16.) Pp. 44. Winnipeg: Verlag der Ukrainischen Freien Akademie der Wissenschaften, 1953.
- The phonemes of English: A phonemic study of the vowels and consonants of standard English. By A. Cohen. Pp. [viii], 127. The Hague: Martinus Nijhoff, 1952.
- Le pluriel dans les locutions adverbiales de temps et de lieu en slave. By Józef Trypućko. (Uppsala Universitets Årsskrift 1952:7.) Pp. 146. Uppsala: A.-B. Lundequistska Bokhandeln; Wiesbaden: Otto Harrassowitz, 1952.
- Povezanost jezičnih elemenata: Problemi ljudskog izraza. By Petar Guberina. (Matica Hrvatska: Mala knjižnica I.) Pp. 433. Zagreb: Matica Hrvatska, 1952.
- Problem 'slaganja vremena' [Le problème de la concordance des temps]. By Guberina Petar [i.e. Petar Guberina]. (Jugoslavenska Akademija Znanosti i Umjetnosti.) Pp. [45]. Zagreb, 1951.
- Prolegomena to a theory of language. By Louis Hjelmslev, translated by Francis J. Whit-

- field. (Indiana University publications in anthropology and linguistics, Memoir 7 of the International journal of American linguistics: Supplement to IJAL, Vol. 19, No. 1.) Pp. [iv], 92. Baltimore: Indiana University (under the auspices of Linguistic Society of America [and] American Anthropological Association), 1953.
- Quinti Septimi Florentis Tertulliani opera; Pars I. Opera catholica adversus Marcionem. (Corpus Christianorum, Series latina, Vol. 1, [published under the direction of the Abbey of St. Pierre de Steenbrugge].) Pp. xvii, 75, with 4 folding charts. Turnholti [= Turnhout, Belgium]: Typographi Brepols Editores Pontificii, 1953.
- Reflections on the numerals 'one' and 'two' in ancient Indo-European languages. By J. Gonda. Pp. 80, offset. Utrecht: A Oosthoek, 1953.
- Revue des langues romanes 71.159-246 (1952).
- Slovo 1.1-86 (1952). Zagreb: Časopis Staroslavenskog Instituta.
- Smithsonian Institution, Bureau of American Ethnology, Bull. 145: The Indian tribes of North America. By John R. Swanton. Pp. vi, 726, with 4 folding maps. Washington: U. S. Government Printing Office, 1952.
- Smithsonian Institution: Sixty-ninth annual report of the Bureau of American Ethnology, 1951-1952. [By Matthew W. Stirling.] Pp. 30. Washington: U. S. Government Printing Office, 1953.
- Speculum 28.1-252 (1953).
- Die Sprache des Amerikaners: Eine Einführung in die Hauptunterschiede zwischen amerikanischen und britischem Englisch der Gegenwart. By Hans Galinsky. Vol. 2, Wortschatz und Wortbildung—Syntax und Flexion, pp. x, 522. Heidelberg: F. H. Kerle Verlag, 1952.
- Strój krzeczowski. By Janusz Świeży. (Atlas polskich strojów ludowych; Część 5: Małopolska, Zeszyt 7.) Pp. 48, with 6 plates. Poznań: Polskie Towarzystwo Ludoznawcze, 1952.
- Strój kurpiowski puszczy białej. By Maria Żywirska. (Atlas polskich strojów ludowych; Część 4: Mazowsze, Zeszyt 5.) Pp. 54, with 7 plates. Poznań: Polskie Towarzystwo Ludoznawcze, 1952.
- Studien zur indogermanischen Grundsprache. Ed. by Wilhelm Brandenstein. (Arbeiten aus dem Institut für allgemeine und vergleichende Sprachwissenschaft, Heft 4.) Pp. 75, offset. Wien: Gerold & Co., 1952.
- A survey of verb forms in the eastern United States. By E. Bagby Atwood. (Studies in American English, No. 2.) Pp. viii, 53, with 31 maps. [Ann Arbor]: University of Michigan Press, 1953.
- Le système consonantique du breton: Avec une étude comparative de phonétique expérimentale. By F. Falc'hun. (Ouvrage publié avec le concours du Centre national de la Recherche scientifique.) Pp. 194. Rennes: Plihon, Libraire, 1951.
- Le système verbal dans les inscriptions 'royales' présargoniques de Lagaš: Contribution à la grammaire sumérienne. By Edmond Sollberger. (Thèse présentée à la Faculté des Lettres de L'Université de Genève ..., Thèse No. 146.) Pp. xvi, 263. Genève: Librairie E. Droz, 1952.
- Die tasmanischen Sprachen: Quellen, Gruppierungen, Grammatik, Wörterbücher. By Wilhelm Schmidt. (Comité International de Linguistes: Publications de la Commission d'Enquête Linguistique. Mit unterstützung der ... UNESCO, etc.) Pp. 521. Utrecht-Anvers: Uitgeverij Het Spectrum, 1952.
- Termini juridici. By Jens Bjelke, ed. by Sigurd Kolsrud. (Norsk historisk Kjeldeskrift-Institutt.) Pp. [xiv], 158, offset. Oslo: Jacob Dybwad, 1952.
- Tijdschrift voor nederlandse taal- en letterkunde 70.161-240 (1952).
- The triumph of the English language: A survey of opinions concerning the vernacular from the introduction of printing to the Restoration. By Richard Foster Jones. Pp. xiii, 340. Stanford, Cal.: Stanford University Press, 1953.
- Two Old Portuguese versions of the Life of Saint Alexis: Codices Alcôbacenses 36 and 266. By Joseph H. D. Allen Jr. (Illinois Studies in Language and Literature, Vol. 37, No. 1.) Pp. x, 67, with [21] plates. Urbana: University of Illinois Press, 1953.

Valeur logique et valeur stylistique des propositions complexes en français et en croate.
By Petar Guberina. Pp. 350, Zagreb, 1939.

Virittaja 56.241-356 (1952).

Westtocharische Grammatik; Band 1. Das Verbum. By Wolfgang Krause. (Indogermanische Bibliothek, 1. Reihe: Lehr- und Handbücher.) Pp. xvi, 312. Heidelberg: Carl Winter, Universitätsverlag, 1952.

Zeitschrift für vergleichende Sprachforschung 71.1-128 (1953).

Zestig jaren onderwijs en wetenschap aan de Faculteit van de Wijsbegeerte en Letteren der Rijksuniversiteit te Gent. (Rijksuniversiteit te Gent: Werken uitgegeven door de Faculteit van de Wijsbegeerte en Letteren, 114^e aflevering.) Pp. 212. Brugge (België): De Tempel, 1952.

ZH guide: An introduction to Sinology. By George A. Kennedy. (Yale University, Sinological Seminar.) Pp. [viii], 171, offset. New Haven: Yale University, Far Eastern Publications, 1953.

Zur Inversion in deutschen Satawörtern. By Sten Hagström. (Uppsala Universitets Årsskrift 1952:8.) Pp. 90. Uppsala: A.-B. Lundequistska Bokhandeln; Weisbaden: Otto Harrassowitz, 1952.

Zvuk i pokret u jeziju. By Petar Guberina. (Matica Hrvatska: Mala knjižnica I.) Pp. 222. Zagreb: Matica Hrvatska, 1952.

31 March 1953

THIS NUMBER OF LANGUAGE

is dedicated

to

FRANKLIN EDGERTON

Sterling Professor Emeritus of Sanskrit

and Comparative Philology

in Yale University

President

of the Linguistic Society of America

in 1934

The publication of this number of *LANGUAGE*
was aided by a grant from Yale University,
which is hereby gratefully acknowledged.

PRÂKRIT *cia*, LATIN *quidem*

JULES BLOCH

Collège de France

La valeur emphatique de l'indo-européen **k^wid* est attestée par *cid* dans le Rgveda,¹ par *ci* de l'Avesta,² par *ciy* du vieux perse,³ par *oûki* d'Homère et d'Hérodote,⁴ par *cid* en vieil irlandais.⁵ La nuance indéfinie de ce petit mot à l'origine, restée visible en présence d'un interrogatif ou dans le cas du redoublement, s'est ailleurs modifiée, passant de 'en quelque chose' à 'un peu', d'où par litote 'plutôt, bien sûr, même, en particulier' (Bartholomae, *Alt. Wb.* 593 'und zwar'). Dans RV IV.3.4 Macdonell lui attribue un sens atténué, Geldner au contraire met le mot en valeur: *tvaṃ cin naḥ śamyā* (pada *śamyai*) *bodhi ... svādhīh* 'merke du wenigstens auf diesen Opferdienst von mir'. Et pour I.10.9 *nū cid dadhiṣva me gīrah*, Macdonell traduit: 'Even now take to thyself my songs', là où rien ne légitime le sens de 'maintenant encore'; mais Geldner écrit: 'Nimm doch ja meine Lobreden an'.

Appelant au sacrifice Indra et Agni, le récitant de VII.93.6 s'autorise des occasions précédentes en disant: *nū cid dhī parimamnāthe asmān*, ce qu'on peut traduire par 'jamais en effet vous ne nous avez dédaignés'; mais de façon plus intéressante, et peut-être plus exacte, par 'du moins jusqu'à présent'.

Nous sommes ici tout près des valeurs du latin *quidem*, ce qui n'a rien d'étonnant, puisque *quidem*, d'où qu'en vienne la finale, repose sur *quid*. On peut dès lors se demander si la forme élargie n'existe pas, elle aussi, sur le terrain indien. Elle existe, mais pas en sanskrit; ce ne sera pas, on le sait de reste, le seul cas où le moyen-indien aura conservé des archaïsmes inconnus du sanskrit et même de l'iranien. Or, qui osera construire par hypothèse une forme sanskrite correspondant à *quidem* aura le choix entre **cidam* et **cida*: c'est le second qu'on retrouve en prâkrit sous la forme *cia* équivalent de *eva*, avec sens restrictif, *avadhārane*.⁶ On note couramment *ccia*, preuve que la particule, pour être enclitique de position, n'est pas traitée phonétiquement comme telle: c'est le signe de sa force expressive.

Les sens les plus courants en sont 'précisément, même'; et avec une valeur adversative qui la rapproche de *quidem*, 'seulement'. Or il y a des cas, rares il est vrai, où il serait difficile de traduire par l'un de ces mots: Hāla, Sattasāi 88,

*jāñai jāñāveuṃ aṇuṇaavidāyamāṇaparisesaṃ
pairikkammi vi viṇāvalambanaṃ sa ccia kuṇantī*

'Elle sait lui faire comprendre que les câlineries n'ont pas dissipé toute sa colère, / en restant, EN CE QUI LA CONCERNE, fidèle à la décence même quand elle est seule'; ib. 93,

¹ Delbrück, *Altindische Syntax* §243; Macdonell, *Vedic grammar for students* 280.

² Bartholomae, *Altiranisches Wörterbuch* 588; Reichelt, *Avestisches Elementarbuch* 444.

³ Kent, *Old Persian* 184; cf. Darius I: *yaθā paruvam ciy* 'comme déjà auparavant' répété trois fois, en regard de *tyaiy paruvam xšāyaθiyā* 'ils ont régné auparavant'.

⁴ Meillet-Vendryes, *Grammaire comparée des langues classiques* §880.

⁵ Vendryes, *L'évolution de l'adverbe cid en vieil irlandais* 4.

⁶ Références chez Luigia Nitti, *Les grammairiens prâkrits* 205, *Prâkṛtānuśāsana de Puruṣottama* 72.

*avianhapeccanijjena takkhaṇam, māmi, teṇa diṭṭheṇa
siṇaapiṇeṇa va pāṇiṇeṇa taṇha ccia ṇa phittā*

'Spectacle dont je ne peux me rassasier, à cet instant, tantine, je l'ai vu; / (mais c'était) comme de l'eau bue en rêve: ma soif, ELLE, n'en a pas diminué'; cf. encore 183, 215.

Le prākṛit écrit souvent *cea* avec la même valeur que *cia*. Les commentateurs interprètent les deux formes par skr. *caiva*, ce qui ne peut se considérer comme une explication phonétiquement valable: il n'y a pas de *ea* issu de *eva* ni de *ia* issu de *iva*. Par contre on peut se demander si le sens adversatif de *caiva* qui s'est développé en sanskrit ne résulte pas du contre-coup du *ccea*-*ccia* usuel, lequel ne pouvait se noter avec quelque exactitude dans la langue savante.

La particule a continué de vivre, et en particulier sur le terrain qui était celui du prākṛit type: le vieux marathe a *-ci*, enclitique emphatique, réduit à *-c* dans la langue moderne. Contigu au marathe vers l'Est, le chatisgarhi⁷ compte parmi ses particules d'insistance *c*, *ec*, 'which mean "even": *dai-c kā* even to the mother; *morec* even mine, *tōroc* thy also'. Le *Linguistic survey of India: Eastern Hindi* donne dans un specimen de Jashpur, donc aux confins du domaine bihari, *toroc ṭhan* 'à ton égard aussi' (216) à côté de *torec* 'thine even' (217); et de même *sab-ec din* 'toujours' etc. Le même élargissement a été noté aussi dans le bhunjia de Raipur, dialecte à vrai dire assez incorrect du halbi, parlé à côté du chatisgarhi et de l'oriya (*LSI: Marathi*): *tūco* 'you', *hunā co* 'him' (377).⁸

Un tel élargissement est susceptible de faiblesse phonétique: on ne s'étonnera pas de trouver loin à l'Est, chez les Cakma de la région montagneuse de Chittagong (*LSI: Bengali* 329, 331), *tūs pele ya* 'husks having-got even', *kai-ya* 'near also'. C'est de façon analogue que doit s'expliquer déjà *ṇai* du prākṛit (Hemacandra II 184) et de l'apabhraṃśa (Pischel, *Materialien* 63, XXXIV = Alsdorf, *Ap. Studien* 104): ce *ṇai* ne représente pas, comme le veulent les commentateurs, skr. *nāpi* mais **na-ci(d)*, cf. grec *οἷκι*.

L'apabhraṃśa a noté le stade intermédiaire *ji*, *je*, également synonyme de skr. *eva*, et assez pauvre d'expressivité dans les textes, à ce qu'il semble; Hemacandra II 217 le compte dans les particules explétives. Mais certaines langues modernes, cette fois-ci parmi celles de l'Ouest et du Centre, témoignent que la valeur d'insistance a persisté. La forme courante est *-j* en guzrati et en sindhi; le marwari (*LSI: Rajasthani* 30) a *-ṭj* ou *hāj* que Grierson rapprochait de skr. *jjeva*, successeur de *yjeva*; il donne comme exemples *in sū hāj* 'even from these', *Mārwar to phāydo huwai ṭj* 'Marwar itself will certainly profit'; il y a même une forme redoublée: *karsī jej* 'he will certainly do'. Enfin on trouve en vieux braj *ju* (non noté cependant par Dharendra Varma).

Etant donné la dispersion de ces diverses formes d'origine commune, on peut imaginer qu'à l'époque moyenne *cia*, et plus tard ses dérivés, ont été plus répandus dans l'Inde aryenne que les témoignages ne le révèlent; il n'est pas interdit de penser qu'on y retrouvera encore de nouvelles traces d'une particule, que son caractère expressif a exposée au renouvellement dans tout l'indo-européen sauf dans les langues marginales extrêmes.

⁷ Hiralal, *Grammar of the Chhatisgarhi dialect of Hindi* 117.

⁸ Ce serait une erreur de chercher la trace de *cia* dans l'enclitique *ce*, *cai* qui en dogra de Jammu (*LSI: Western Hindi and Panjabi* 643) 'is added to the imperative to give a permissive force. Thus *khācai* let us eat, *manācai* let us celebrate'. Il s'agit d'un équivalent (réservé à la 1^{re} personne?) de hindi *cāhe* exprimant le souhait, qu'on retrouve avec la même valeur 'permissive' dans nepali *cāhī*, *cāt* et assamais *con* (B. Kakati, *Assamese* §824).

OBSERVATIONS SUR LES COMPOSÉS NOMINAUX DU RGVEDA

LOUIS RENOU

Sorbonne

1. Par rapport à l'usage du sanskrit classique (très variable d'ailleurs suivant les genres et les époques), le Rgveda (RV) ne présente qu'un nombre relativement limité de composés. Laissant à part les formations nombreuses, mais peu caractéristiques, où le membre antérieur est un invariant—préverbe, préposition, préfixe, adverbe ou nom adverbialisé—et qui n'entrent pas dans la définition stricte de 'composé nominal', on ne trouve en moyenne que cinq composés pour 150 mots (ou: pour 80 mots pleins), à peine davantage dans les portions récentes de la Samhitā. La proportion varie suivant le style: les passages descriptifs (ainsi 2.21) ou emphatiques en présentent davantage, les hymnes simples (comme l'hymne à Pūṣan 6.54 ou le dialogue 'affectif' Yama-Yamī 10.10) en ont fort peu.¹

Les catégories a priori possibles, en fait largement développées pour la plupart en sanskrit ultérieur, sont loin d'être exploitées en totalité.

2. Prenons le cas de la combinaison la plus simple, les dvandva. Seul le type *mitrāvāruṇā* (non séparé dans le texte pada!)—c'est-à-dire les devatādvandva—est bien représenté, avec les variantes morphologiques qui progressivement en dérivent.² Il n'existe que trois exemples isolés de dvandva pluriel: l'un, *ahorātrāṇi* du Livre X,³ est refait sur le duel, le second, le voc. *indrāmarutaḥ*, comporte un pluriel tantum au membre ultérieur, le troisième, *ukthārkā* (désinence neutre insolite) appartient à une petite série de juxtaposés d'inspiration rituelle, imités des devatādvandva.⁴ L'unique dvandva singulier (neutre) est *iṣṭāpūrtā*, du Livre X; Wackernagel, *Ai. Gr.* 2.1.160.⁵

Il n'y a, pareillement, qu'un seul dvandva adjectif sûr, fait d'ailleurs d'éléments substantifiés: *nīlalahitā* du Livre X 'chose qui est à la fois (ou: mixte entre) bleu et rouge'.⁶

Il est donc difficile d'imaginer une formation plus unilinéaire. Le succès des devatādvandva a été dû à la fréquence des invocations jointes à deux divinités (encore que l'invocation jointe à plus de deux n'ait pas entraîné de dvandva trimembre); un facteur déterminant a été l'effective coopération qui d'ordinaire accompagne une telle jonction sur le plan ritualiste ou mythologique.⁷

¹ Les composés à trois membres sont extrêmement rares, Macdonell, *Vedic gr.* §241, et (plus complet) Wüst, *Stilgeschichte* 56. Ils répondent presque tous à la formule (1 + 1) + 1 ou 1 + (1 + 1). Dans *rtājātasatya* et *vakmarājasatya*, l'élément *satya* est un appendice semi-explétif, préluant à l'appendice 'sat du sanskrit tardif, plus spécialement du bouddhique, *BSOS* 9.48.

² Cf. Wackernagel, *Ai. Gr.* 2.1.150-6.

³ Cf. Thieme, *Jha commemoration* vol. 416.

⁴ Le prototype étant *ukthāmadā* (AV) = *ukthām ca mādaś ca* Hillebrandt, *BB* 9.192.

⁵ Il est vrai qu'il faudrait ajouter les dvandva numériques où le membre ultérieur est un nom de dizaine, de centaine ou de millier, qui demeure naturellement au singulier; mais ces numériques sont eux-mêmes fort peu abondants dans le RV; Wackernagel 3.379, 382.

⁶ Autres exemples chez Wüst 59, tous douteux.

⁷ Pas de dérivés de devatādvandva (qui abondent en revanche à l'époque des Brāhmaṇa), sauf *mitrāvāruṇavānt* (hapax) et *maitrāvāruṇā*, semi-patronymique.

3. La situation dans les tatpuruṣa est évidemment plus complexe. La catégorie dominante est celle des tatpuruṣa verbaux, c'est-à-dire où la relation d'un membre à l'autre s'exprimerait en phrase libre par la relation d'un verbe avec son régime, notamment un régime direct (accusatif). Dans la série des noms d'agent, on a ainsi le type *haviṛād* 'qui mange l'oblation' ou *viśvajinvā* (voc.) 'qui incitent toutes choses'.⁸ A *viśvajinvā* répond exactement le type, plus rare, *radāvaso* (voc.) 'qui dispense la richesse', c'est-à-dire avec ordre inversé des membres. En général le type *radāvaso* se présente modifié, normalisé, en **radad-vaso*, par exemple *vidādvasu* 'qui gagne la richesse': ceci a été possible du jour où un élément tel que *rada*, faussement senti comme impératif, a pu aisément se transformer en **radat*, c'est-à-dire en injonctif à nuance impérative.⁹

Un renversement de membres analogue à *viśvajinvā/radāvaso* se présente (a) dans le groupe fort restreint, mais ancien (Wackernagel 320)¹⁰ de tatpuruṣa verbaux, du type *dātivāra* 'qui donne des trésors', en regard de *jarādaṣṭi* 'qui atteint la vieillesse' et (moins nets) *vasudhiti*, *vasunūti*, *havyādāti*; (b) dans le groupe plus nombreux où l'élément verbal est représenté par l'adjectif en *-ta*, type *sulāsoma* 'qui a pressé le soma' (liste des formes chez Wüst, *Stilgeschichte* 66): mais ici la forme inversée, type *putrahata*, n'apparaît qu'après le RV, Wackernagel 302.¹¹ Il est très probable que plusieurs des formes précitées en *-ti* ont été versées secondairement dans la catégorie en *-ta* (Wackernagel 276), cf. *vīdhavya* à côté de *vīthotra*.

4. En regard de cette masse considérable de tatpuruṣa verbaux, les tatpuruṣa nominaux sont relativement rares. On pouvait cependant s'attendre que la jonction syntaxique entre deux noms, soit apposés, soit régimes l'un de l'autre, donnât lieu à une considérable fabrication de composés; il n'en est rien. Certes, les composés déterminatifs en °*pati* sont fréquents, mais c'est sans doute parce que justement le mot *pāti* y était senti moins comme un membre autonome de composé que comme une manière de suffixe, analogue à *-van-* ou *-vant*,¹² de même qu'il est arrivé dans les finales en °*dā*, en °*vasu* et dans bien d'autres. Hors de ces formations en °*pati*, il n'existe, du moins dans le RV ancien, qu'un petit nombre de tatpuruṣa déterminatifs, des mots isolés comme *rājaputrā* et *brahmaputrā* (°*putrā* étant fait d'après °*pati*), *drupadā* (où °*padā* est également

⁸ La valeur passive, assez fréquente (Wackernagel 175), du nom-racine, type en °*yūj* etc., atteste un renforcement de la valeur 'verbale'. De même une expression comme *havanāsrūt ... bodhi* 2.33,15 'sois (tel que) tu écoutes notre appel' = 'écoute notre appel'.

⁹ Il s'y ajoute le fait qu'une forme comme *tarād[d]veṣas* 'trionphant des ennemis' était résoluble à volonté en *tara-dveṣas* (type ancien) ou en *tarad-dveṣas* (type nouveau), cf. BSL 41.217.

¹⁰ Ajouter les noms propres *vr̥ṣṭihavya* (ton!) (cf. *vr̥ṣṭidyū*) et *śr̥ṣṭigu* (cf. *pūṣṭigu*), littéralement 'qui fait pleuvoir l'offrande' et 'qui fait obéir les vaches'.

¹¹ Les composés terminés par un verbal en *-ta* ne connaissent pas la valeur accusative du membre antérieur; toutefois *stómataṣṭa* 10.15,9 est sur la voie, le contexte invitant à rendre 'ceux qui ont façonné les panégyriques (par leurs chants)', ailleurs, en revanche, '(poésie) façonnée en forme de panégyrique'.

¹² De là le 'samdhi interne' de *viśpāti* comme *viśva*, et le type redondant si fréquent *rayīṇām rayipātiḥ*, ou encore *sātpātiḥ* (= *sattamam*) *pātim* 1.11,1. °*Pati* au sens du suffixe *-vant-* est connu du sanskrit ultérieur, cf. la juste observation de H. W. Bailey, BSOAS 14.421.

senti comme un suffixe), *devajand* (= *devātāt*, même remarque), *aśvayūpā*, *hiraṇyapiṇḍā*, *udameghā*. En outre, quelques noms propres où le ton antérieur révèle apparemment une influence venue des bahuvrīhi: *dīwodāsa*, *tānūnapāt*, *devākṣatra*, *udāvraja*. L'appellatif *bāhvōjas* 'force des bras' (à côté de 'fort des bras', bahuvrīhi) atteste cette même influence.

5. La relation appositionnelle d'un membre à l'autre n'est exprimée que dans deux formations, l'une et l'autre du RV récent et avec ton antérieur (ce qui laisse un doute sur leur authenticité): à savoir, *ūlūkayātu* (et similaires, dans une strophe d'hymne d'Anhang) et le nom propre *vṛṣākapi*.¹³

La classe adjectif (épithète) + substantif ne comprend qu'une seule formation bien accréditée, *candrāmas* 'lune'; en outre, quelques rares composés en *mahā°* ou en *eka°* (auxquels joindre *ardhadevā*), autrement dit, à membre antérieur adverbialisé; enfin, le nom propre du Livre X *saptarṣāyah* et le semicomposé *madhyāmdina*.

Restent les tatpuruṣa dont le membre ultérieur est un adjectif. Là encore il n'y a, comme formes sûres du RV II-IX, que quelques hapax, en grande partie avec maintien d'une désinence interne: *māderaghu*, *gāviṣṭhira* (nom propre) et *jātūṣṭhira*, *yajñādhīra*, *tanūśubhra*, sans compter plusieurs mots où l'élément antérieur équivaut à un adverbe ou à un préfixe.

Bref, rien ne laisse prévoir dans le RV ancien l'extension progressive et rapide dont bénéficieront les tatpuruṣa nominaux en sanskrit ultérieur: c'est là le point décisif où le procédé compositionnel submergera de proche en proche la plupart des relations syntaxiques normales et fera échec aux habitudes flexionnelles si solidement établies dans le Veda. Les tatpuruṣa nominaux du RV, en partie hérités d'ailleurs (ceux en *°pati* et sans doute *candrāmas*), sont isolés en général et peu féconds.¹⁴ Il est sensible que, dvandva's mis à part, le RV ne connaît pratiquement que deux formations, les tatpuruṣa verbaux pour l'expression nominale abrégée, rapide, de relations à caractère verbal; et, d'autre part, les bahuvrīhi, pour l'expression de relations à caractère proprement nominal, les relations mêmes que rend, en phrase libre, une proposition relative ou circonstancielle à valeur descriptive.

6. Les bahuvrīhi en effet servent d'épithètes ornementales, pour décrire l'activité ou l'attitude des dieux sous forme concrète, imagée; nombre d'entre eux impliquent une comparaison; un mot comme *saptāśya* 4.51,4 suffit à évoquer le nom du dieu (Bṛhaspati). Les bahuvrīhi, étant essentiellement nominaux, ne participent pas au dynamisme de la phrase; on peut les supprimer sans que rien manque au procès. Le fait que le membre antérieur soit, le cas échéant, un participe comme *dadṛṣāṇḍpavi* ou *hatāvṛṣṇāḥ* ne change rien à la nature du composé. Il n'existe pas de bahuvrīhi 'verbal',¹⁵ une fois qu'on a pris soin de dissocier le

¹³ *Vṛṣanāma* 9.97,54 est à lire *vṛṣa nāma* (Geldner ad loc., qui compare justement *nāma dhenū* 6.66,1).

¹⁴ On dit *sahasah putrāḥ* ou *sūnūḥ*, mais (en tatpuruṣa 'verbal') *sāhaskṛtāḥ* ou *sahojā*, comme désignation d'Agni. Ou encore *hiraṇyarūpāḥ*, *hiraṇyavarṇāḥ* etc. (bahuvrīhi's), ou bien *hiraṇyadā* (tatpuruṣa verbal), mais *hiraṇyāyād yōneḥ* 2.35,10 (pas de tatpuruṣa nominal!). L'officiant est *yajñāketuḥ* 4.51,11, mais l'Aurore est *yajñāśya ketuḥ* 1.113,19.

¹⁵ *Tvāṃkāma*, l'unique apparent bahuvrīhi 'verbal', est une extension de la catégorie *dhanamjayā* (§8).

groupe d'avec les formes à membre antérieur 'régissant' (ci-dessus §3), qui ne sont que des tatpuruṣa verbaux inversés; on a eu tort de les considérer comme des bahuvrīhi aberrants.

Les bahuvrīhi du RV présentent déjà tous les caractères de ceux de l'époque classique (sauf en ce qui concerne le nombre des membres, strictement limité à deux): notamment la liberté dans les rapports d'un membre à l'autre, ou l'éventuelle appartenance du membre antérieur à un mot situé au dehors; v. Wackernagel 33.

7. Pourtant, (a) les séries classiques productives, de type classificatoire, en °ādi, °rūpa, etc., sont à peine amorcées. Peut-être faut-il avec Geldner en voir l'ébauche dans les formes du RV récent *dāśāpravarga* et *dāvabudhya* 1.92,8, *dāvabudhya* et *gāgra* ib. 7 (*dāvabudhna* aussi au Livre X); cf. encore *īndrajyeṣṭha* et *samudrājyeṣṭha*. (b) Les bahuvrīhi substantifiés ne sont attestés qu'après préfixe ou élément assimilable à un préfixe (Wackernagel 304); ou encore dans le sous-groupe des dvigu (Wackernagel 305), dont les rares exemples ṛgvédiques sont du Livre X: *triyaṅgā*, *daśāṅgulā*, *daśāntaruṣyā* (*tryudāyā* obscur, Geldner).

8. On sait qu'une grande quantité de composés maintiennent la désinence du membre antérieur. Le fait se présente variablement: fort peu dans les bahuvrīhi; peu aussi dans les tatpuruṣa nominaux, sauf dans ceux en °pati chez lesquels la finale -as- est fréquente, qu'elle soit ou non authentiquement une désinence de génitif. Au contraire, les tatpuruṣa verbaux présentent fréquemment la désinence interne -m, tout au moins dans la formation quadrisyllabique *dhanamjayā* (liste Wüst 69) où l'insertion a été utilisée rythmiquement. Enfin (sans parler des āmreḍita et autres types inorganiques¹⁶), tous les dvandva de structure ancienne. Mais le maintien de la désinence repose sur des conditions diverses. Dans le cas des dvandva, on a reconnu depuis longtemps (quoique Wackernagel 151 ne se prononce pas) que la source de la formation était le duel isolé, dit à tort 'elliptique',¹⁷ celui qu'on trouve dans *mitrā* 'Varuṇa et Mitra' ou dans *vāruṇa* (-a bref!)¹⁸ 'id.' 5.64,6; 7.61,1 et 64,5; la jonction pleine *mitrā-varuṇā* est secondaire, comme semble l'indiquer la répartition textuelle. Il existe du reste aussi un pluriel 'elliptique', par exemple dans *mitrāsah* 7.38,4 'Varuṇa, Mitra, Aryaman' (mécompris, puisqu'il coexiste avec *vāruṇah* et *aryamā* au sing.) et dans *aryamānah* 5.54,8 'id.' (Geldner).

Toute autre est l'origine du type *dhanamjayā*: il ne repose pas sur une juxtaposition directe *dhāna* + **jaya*, mais sur la formule verbale *dhānam jayati* ou *jayatu* ou *jaya*; le suffixe note la fonction verbale en composition. La remarque vaut, au surplus, pour tous les tatpuruṣa verbaux, qu'ils aient ou non une désinence médiane. Ainsi *gopīthā* 'acte de garder les vaches' n'est pas *gō* + **pīthā*, mais (*go* + *pī*)*thā*; *sōmapīti* n'est pas *sōma* + *pīti*, bien que *pīti* existe aussi et qu'un groupement *sōmasya pīti* soit secondairement attesté.

9. Le style védique comporte dans une large mesure (et bien plus que les

¹⁶ Wackernagel 325; Wüst 73; K. Hoffmann, *IF* 60.254. Ajouter *kenipā*, s'il est permis, (avec Velankar) d'analyser ce terme obscur en *ké nīpānti*, et de même *kēpi* en *ké pipyuh*: *Journ. Univ. Bombay* 21.2.15-6 (1952); cf. *katpayā*.

¹⁷ Le duel 'elliptique' n'est pas limité aux devatādvandva, cf. *āndhasī* = Soma et Surā (Geldner ad 7.96,2) et peut-être *rōdasī*.

¹⁸ Du moins au vocatif (duel), Wackernagel 53 §20a.

traductions ne le laissent voir) la juxtaposition (asyndète) de deux substantifs, qui selon nos usages seraient reliés par 'et' ou même subordonnés l'un à l'autre.¹⁹ Le type *sīndhur nā kṣōdah* 'comme un fleuve, (comme son) flot' a été depuis longtemps identifié par Bergaigne, *Mélanges Renier* 93; on dit indifféremment *gāvām nā sārgāh* 4.51,8 'comme des circulations de vaches', ou *gāvo nā yūthām* 8.46,30 'comme des vaches, (comme leur) troupeau'. De tels emplois n'expliquent pas directement la naissance de composés, puisque, on l'a vu, la jonction de deux substantifs est rare, du moins en tatpuruṣa. On ne peut obtenir de composé en partant de *sōme mādē* 1.80,1 'dans l'ivresse (du) soma', encore moins, de *pārvato gīrth* 1.37,7 'le rocher, la montagne'; le type *vrtrāya hāntave* n'a pas davantage abouti à **vrtrahantave*. Mais ces juxtapositions sont tout de même le préliminaire aux composés. Ce n'est pas un hasard si le rapport interne des membres dans le bahuvrīhi implique une proposition comparative si souvent, celle même que dégage l'expression analytique *sīndhur nā kṣōdah*. Plusieurs juxtaposés asyndétiques sont aptes à s'accoler en bahuvrīhi, ainsi *aghāśamsa* 'dont la parole est le mal' en face de *aghāt...śāmsāt* 1.166,8; même en tatpuruṣa, *sūktavākā* (Livre X) en face de *sūktāya vācase* 9.90,6; *devān jānma* 'la génération (, celle des) dieux' 1.71,3; 6.11,3 est à mi-chemin entre l'attendu *devānām jānma* et un **devajanma*, fait comme *devajanā*. La langue hésite entre *itā ūt* et *itāūti*, *itthā dhtyaḥ* 6.62,3 et *itthādhiye* 4.11,3. Déjà Wackernagel 288 avait relevé quelques exemples de 'pré-bahuvrīhi', semblant refléter un état où le procédé compositionnel était encore mal fixé. On peut ajouter à cette brève liste des formes comme *dīnā dākṣāh* 4.24,9 'eux, aux faibles capacités', *ēkam krātuḥ* 6.9,5 = **ekakratum*, *vīrenyaḥ krātuḥ* 10.104,10 = **vīrenyakratuḥ*, cf. Geldner ad loc., *bhadrāḥ krātuḥ* 1.91,5 = **bhadrakratuḥ* (analogues 4.10,1; 1.107,3), *brhān kṣāyaḥ* 8.15,9 = **brhathkṣayaḥ*, cf. Geldner; *sāho mahāt* 5.11,6 équivaut à **mahāsahāh*. L'instrumental descriptif, qui n'a pas de vitalité en védique, masque un état pré-compositionnel, ainsi *sthīrēbhīr āngaiḥ* 2.33,9 = **sthīrāṅgaḥ*, *samānēbhīḥ paṁsyebhīḥ* 1.165,7 = **samānapaṁsyebhīḥ*. Souvent ces groupes analytiques sont compris comme l'expression d'une parenthèse ou d'une anacoluthie, ainsi *aruṣāstūpo rūśad asya pājāḥ* 3.29,3 'avec son aigrette rouge—blanche (en revanche était) sa masse corporelle'; de même (presque toujours avec *asya*) 1.115,5; 163 1 et 9; 186,8; 5.62,7; 6.75,11 et 15.²⁰

10. À côté de ces emplois pré-compositionnels, il existe aussi des composés scindés par tmèse. Aux cas, rares d'ailleurs, groupés par Wackernagel 30 on

¹⁹ Exemples *New Ind. ant.* 3.266 (article auquel nous avons emprunté quelques-uns des faits mentionnés ci-après, pour les présenter dans un autre développement). Cf. aussi Thieme, *Untersuchungen* 32 et suiv., qui cite notamment des substantifs en *-ti-* (*arati*) etc., faussement considérés comme adjectifs par les traducteurs modernes; la même observation vaut pour *rtā*, *krātu*, *sūktā*, *śūṣma* et bien d'autres. Cf. aussi G. Liebert, *Nominalsuffix -ti-*, notamment 142 et suivantes.

²⁰ Dans *barkhīr u tistirāṇā* 1.108,4 on a une forme analytique que remplacera bientôt le composé, autrement structuré, *stīrṇābarkhī*. — A vrai dire, le composé est encore tout proche, dans le RV, de l'énoncé analytique, sans qu'on soit autorisé, naturellement, à parler pour autant de formes 'pré-compositionnelles' chaque fois qu'on rencontre cette variation: *dhūmā te ketūr abhava* 5.11,3 en face de *dhūmāketu*; *yāsya prātīkam āhutaḥ ghr̥tēna* 7.8,1 en face de *ghr̥tāprātīka*, *ghr̥tām ānnam asya* 2.35,11 en face de *ghr̥tānna*; ensemble: *dhūmāketum* ... *yajñānām ketūm* 8.44,10.

ajoutera maintenant *droghāya cid vācase* 6.62,9 (cf. *droghavāc*) et *rjūr ic chāmsaḥ* 2.26,1 (mais l'insertion de *id* n'a pas empêché le composé dans *tadīdartha*), où le membre antérieur a été adapté à sa fonction grammaticale, alors que dans *nārā ca (vā) śāmsam* il est resté intact.²¹

Sans doute faut-il voir un composé secondairement défait dans *śitām gābhastim aśānim* 1.54,4 (Geldner) et dans une série de noms propres, *pajrāya sāmne* 8.4,17 et 6.47, *rūsato vāpsasaḥ* 1.181,8, *jārataḥ kārṇam* 10.80,3, *varo suśāmne* 8.23,28; 24,28; 26,2 (?), qui préludent aux fantaisies du type *daśapūrvaratha* en sanskrit classique poétique.²²

Le groupe *mānaso jāveṣu* 10.71,8 'dans les improvisations de la pensée' remplace assez complexement un bahuvrīhi **manojavaṣu*, épithète de *stōmeṣu* à suppléer (noter que le comparatif de *manojū* ne pouvait être autre que *mānaso jāvīyān*). *Divī kṣāyaḥ* 3.2,13; 8.64,4; 10.63,5 représente la scission d'un **divikṣaya*, attesté ailleurs. Dans *divo nā sādmanakhasam* 1.18,9, *divī sādmanabarhiṣaḥ* 1.52,4, *divā ā prṣṭhayājvane* 5.54,1, *śūṣmaṃ nṛṣāhyam* 9.30,8, il s'agissait sans doute d'éviter un composé trimembre.

11. Ce sont là les précurseurs de ce que M. Helmer Smith a appelé le 'split-compound'.²³ Il est intéressant de voir que sur ce point comme sur celui des formations pré-compositionnelles, le sanskrit bouddhique hybride a conservé des habitudes apparemment héritées du Veda: cf. les notes de Weller au Lalitavistara (2.22; 17.13; 429.8 et 13),²⁴ celles de Senart au Mahāvastu I.393, 466, 499, etc. Des emplois comme *mahāpadmo yathodbhavaḥ* ou *gaṅgā yathā vālikā* (cités Senart I.396) sont absolument identiques aux juxtaposés du RV en *sindhur nā kṣōdah*. Le grand ouvrage de M. Edgerton, qui sera paru quand ces lignes s'imprimeront, contribuera peut-être à mettre en évidence d'autres faits de ce genre.²⁵

²¹ Tmèse en pāli: H. Smith, *Critical Pāli dict.* 1.xxvi et 33*. En sanskrit classique, ma *Grammaire* 84 §76 n.

²² Exemples Hultzsch, *South-Ind. inscr.* 1.166 n. 6; Śis. 1.42; R. Schmidt, *Nachträge* sous *arāṭa*, etc.

²³ H. Smith, *Critical Pāli dict.* 1.xxvi et 32*, ainsi que BSL 33.172. Mais *marāṇe kāle*, qui du point de vue pāli est le 'split-compound' d'un **marāṇakāle*, ne fait que continuer l'usage védique attesté ci-dessus par *sōme māde* et tant d'autres expressions asyndétiques.

²⁴ Les résolutions dans le cas d'expressions numériques complexes (ibid. 31,10; 61,19; 84,1 et 7) ont des correspondants en sanskrit brahmanique, Wackernagel 3.391 §198b α n., ma *Grammaire* 385 §268c n.

²⁵ Des abréviations connues de composés sont *mitrā* pour **mitrātithim* et *patayāt* pour **patayātsakham* (Wackernagel 30); il faut y ajouter *saṃvānanā* 8.1,2 pour **saṃvananāṃkaram* et *dśvaṃ gāṃ rathaprām* 8.74,10 pour **āśvapram* et *goprām* (autre, Geldner). De même *jaritṛṇām* 7.66,3 pour **jaritṛpā* et *dakṣa* (*kavikrato*) 3.14,7 pour **kavidakṣa* (Geldner). Le fait va de pair avec les 'finales syncopées'.

SECONDARY DERIVATION FROM SANSKRIT *i*-STEMS

RULON WELLS

Yale University

1. THE PROBLEM. Sanskrit noun stems ending in *u*, when subjected to secondary derivation, typically replace *u* by *av*, whereas, typically, stems ending in *i* simply drop the *i* before a suffix beginning with a vowel or with *y* (rather than replacing the *i* by *ay* or *e* respectively). I have not found any attempt to explain this disparate treatment, which, to judge from the implication or the express statement of our standard reference works,¹ prevails equally in the Rig-Veda and in all later stages of Sanskrit. To point out, as the grammars do, that in this regard *i*-stems² behave like stems in *a*, *ā*, and *ī* is not to explain the disparity but to heighten it.

Examples: (1) *u*-stems: *mānav-ā* 'Mensch' from *mānu* 'Mensch'; *kaurav-yā* (AV) 'descendent of the Kurus' from *kūru* (VS). (2) *i*-stems: *daivāp-ā* (ŚB) 'descendent of Devapi' from *devāpi*; *āditi-yā* 'Sohn der Aditi' from *āditi*. (3) *a*-stems: *ārks-ā* 'Abkömmling des ṛkṣa' from *ṛkṣa*; *sāhadev-yā* 'von Sahadeva entsprossen' from *sahādeva*.

2. THE ANSWER. Since Sanskrit *i*-stems thus behave differently not only from *u*-stems but also from stems in *r*, *m*, and *n*—that is, the Sanskrit reflexes of the other five PIE semivowels, there is a historical puzzle. In offering an explanation of this puzzle (as briefly as possible), I will first set forth the entire account in outline, with little attempt to distinguish between what is sure and what is conjectural; and then, in §4, will take up the detailed discussion of various points.

The heart of the answer is this. Such a derivative as *āvya* 'vom Schafe her-rührend' from *āvi* (m, f) 'Schaf' is in its origin *āvya*, but comes to be reanalyzed, restructured as *āv-ya*. Then there is room for innovations of types **av-a* and (with *vrddhi*) **āv-a* to match types *av-ya* and **āv-ya*.

3. GENERAL STATEMENT. The explanation is here given in ten subsections.

3.1. PIE has substantive and adjective stems in all six semivowels: *i* (alternating with *i̯* and *j̯*, and similarly for the other five), *u*, *r*, *l*, *m*, and *n*. Stems in

¹ Whitney, Brugmann, Wackernagel, Macdonell, etc., which it is unnecessary to identify. From recent literature I cite the following: T. Burrow, Some remarks on the formation of nouns in Sanskrit, *Annals of the Bhandarkar Oriental Research Institute* 32.19-33 (1951, publ. 1952); J. Kurylowicz, Le degré long en indo-iranien, *BSL* 44.42-63 (1948); id., La loi de Brugmann, *BSL* 45.57-60 (1949); id., *L'accentuation des langues indo-européennes* (1952); L. Renou, *Grammaire de la langue védique* (1952).

All words cited are from RV unless the contrary is stated. When first cited, they are followed (except proper names) by a gloss which, if the word is found in RV, is taken (often incompletely) from Grassmann's *Wörterbuch*. A colon (as in *mānavā* : *mānu*) means 'is a secondary derivative from'; a hyphen (as in *mānav-ā*) means as usual that there is thought to be a morph division between stem and suffix at that point.

² By which I mean all stems ending in *i*, whether this constitutes the whole formative (as in *arc-i* (m) 'Strahl') or is merely the last part of the formative (as in *ma-ti* (f) 'Andacht', *bhū-mi* (f) 'Erde'). Similarly for *u*-stems.

l and *m* are too rare to set a pattern, and may be left out of account. All the four other classes of stems are amenable to secondary derivation, as indeed are all classes of primary noun stems, and all four are then subjected to two main varieties of treatment, which I will call the 'shorter' and the 'longer treatment'. In the shorter treatment one finds from *i*-stems *i* or *iḡ* or *ḡ*, etc.; in the longer treatment, *ei* or *eḡ*. Similarly, in the shorter treatment of *u*-stems the stem ends in *u* or *uḡ* or *ḡ*; in the longer treatment, it ends in *eu* or *euḡ*.³

3.2. By the RV period, the shorter treatment has come to be regular or typical for *i*-stems, the longer for *u*-stems. I have no explanation to offer for this disparate development. The longer treatment for *i*-stems is only preserved in isolated relics, such as *hṛdaya* 'Herz' (for **hṛdāya* from **hṛdi* or **hardi*, PIE *ǵhṛdi* or *ǵherdi*); the shorter treatment for *u*-stems, on the other hand, is not at all uncommon.

3.3. Examples to §3.1 and §3.2. All the Indic words cited are RV unless otherwise labeled.

I. Shorter treatment: (1) *u*-stems: Lat. *aevum* (PIE *aiyo*) 'age' : *āyu* (n) 'Lebenskraft'; PIE *uidhuyā* 'widow' (see Ernout-Meillet); *paidvā* 'dem pedú gehörig' : *pedú* (m); *pārśvā* (n) 'Rippengegend' : *pārśu* (f) 'Rippe'; *yādva* 'den Yadu ... angehörig' : *yādu* (m); Thessalonian *ἡστυεύς* 'Stadtbewohner' : Gk. *ἡστυ* (Schwyzer 472). (2) *i*-stems: *pārthyā* 'Nachkomme des prthi' : *prthi*; *kāvya* 'von kavi stammend', *kāvya* (n) 'Weisheit', both from *kavi*. (3) Other: *sakhya* (n) 'Freundschaft' : *sakhāy* (m, f) 'Genosse'; many other examples in Brugmann, *Gdr.*² 2.1.148-66.

II. Longer treatment: (1) *u*-stems: PIE *uidheyā*, RV *vidhāvā* 'Witwe' : *vidhū* 'vereinsamt'; *vasavyā* (n) 'Reichtum' : *vāsu* (n) 'das Gut.' (2) *i*-stems: Gk. *δοτέον* (**dōteion*) : Skt. *āsthi* 'Bein' (*Gdr.*² 2.1.168); *hṛdaya* (n) 'Herz' : **hṛdi*; *ābhogāya* (m or n) 'Nahrung' : *ābhogī* (f) 'Nahrung'. With *ṽddhi*: *ārṣe-yā* 'von einem ṛṣi ... stammend' : *ṛṣi*; *ādīte-yā* (m) 'Sohn der Aditi' : *āditi*; *maūne-ya* (n) 'Zustand eines mūni' : *mūni*. (These last three examples are differently analyzed by Sturtevant, *CP* 36.356-64.)

3.4. Because of the now regular shorter treatment, the *a*-derivative of *arī* (m) 'der Fromme', for example, would be *aryā* 'fromm'.⁴ But by Edgerton's

³ The choice between *i* and *iḡ* and *ḡ*, like the choice between *ei* and *eḡ*, etc., is allophonic and therefore nonphonemic at some period of PIE—see Edgerton, *The Indo-European semi-vowels*, *Lg.* 19.83-124 (1943); but it has ceased to be so in all the historical IE languages. It may well have ceased to be nonphonemic even in late PIE. The present paper does not commit itself on this question.

Presumably at some period of PIE the choice between the shorter and the longer treatment (for instance, to invent an example, between *ǵhṛdīo* and *ǵhṛdeīo*) is also nonphonemic—an ablaut variation (Brugmann, *Gdr.*² 2.1.198). But by late PIE it had become phonemic. Whether it was exploited semantically after becoming phonemic is not clear. For at least one word, the word for 'widow', there is reason to posit both treatments, *uidhuyā* and *uidheyā*, without a difference of meaning (Ernout-Meillet s.v. *uiduus*). Whatever semantic difference may have existed at one time had disappeared by the time of RV.

⁴ P. Thieme, *Der Fremdling im Rgveda, Abhandlungen für die Kunde des Morgenlandes* 23.2.78-9 (1938): 'Gegen ein von *arī* abgeleitetes Adjektivum *aryā* kann man keinerlei Bedenken erheben. *arī* : *aryā* = *āvi* : *āvya*.' See further A. Debrunner, *Zwei altindische Probleme, New Indian antiquary* ES 1.71-4 (1939), esp. 72-4 'Zum Ariernamen'.

'converse of Sievers's law' (Lg. 10.237, 19.87), the *ya*-derivative of *arī* would also be *arya*; the *a*- and the *ya*-derivatives of *āpī* (m) 'der Befreundete' would both be trisyllabic *āpya* (cf. *āpya* (n) 'Freundschaft'), etc. If not in these particular words, then at least in many words the *ya*-derivative would bear the same placement of the accent as the *a*-derivative. Thus the *a*- and the *ya*-derivatives of some *i*-stems would be fully homonymous. (Cf. Kurylowicz, *L'accentuation* 53, 54.)

3.5. Moreover, all or some of these homonyms would also be synonyms. Sanskrit (including Rigvedic) *ya* is indistinguishable in some of its most frequent secondary uses from secondary *a*, just as PIE *ie/yo* is indistinguishable in many of its most frequent secondary uses from PIE *e/o*.

3.6. As a result, at whatever period the practice of using the 'shorter treatment' for *i*-stems became current, there is no contrast between *a*- and *ya*-derivatives of *i*-stems. A word like *aryá* might be felt or might function now as *ary-á*, now as *ar-yá*; or perhaps, say, *aryá* would always be felt as *ary-á* but *āvya* 'vom Schafe herrührend' would always be felt as *āv-ya*; or there are many other possibilities. But in point of fact *aryá*, *āvya*, etc. came to be felt, all of them, as *ya*-derivatives: *ar-yá*, *āv-ya*, etc.

3.7. In some cases this involved a restructuring; e.g. original *āvya* became *āv-ya*. And thereby room was made for a new creation like **av-a*, or, with *vrddhi*, **āv-a*. In this new type of creation the treatment given to *i*-stems is neither the longer nor the shorter, but what we may call the 'curtailed treatment': the *i* is simply dropped.

3.8. This restructuring is not fully explained in the present paper, but certain factors can be discerned. Some of these give a 'push' to the restructuring, others give a 'pull'.

A push would be given by any *i/a* alternation in which the *i* was felt as primary and the *a* as a secondary displacement of the primary *i*. There are at least two classes of instances that might have given such a push.

(a) Sanskrit *i*-stems in which the *i* is from PIE *ə*. An *a*-derivative from these, if formed when the law 'ə before vowel disappears' was operative, would appear by RV times to simply displace the *i* rather than be added on to it. Most of the alleged *i*-stems with *i* from *ə* are dubious, and the whole notion of such stems is out of favor nowadays, but one quite probable case is *hári* 'feuerfarben, goldgelb'. But this source of *i/a* relationship is only mentioned here as a speculative possibility.

(b) *i/a* alternations arising from *i/an* alternations (*akṣi/akṣān* 'Auge', *ásthi* (AV) / *asthān* 'Knochen', etc.), in which, therefore, the *a* has arisen by regular phonetic change from preconsonantal *ṇ*, and has then been generalized to other positions.

(c) There is another possibility that ought to be mentioned. A proportion such as *vārya* : *vāra* = *pārthyá* : **pārthá*,⁵ in which the relationship is not primarily semantic but merely phonetic, may well have been operative to create a vacuum for such an innovation as **pārtha*.

⁵ With *vārya* (n) 'Kostbarkeit, Schatz'; *vāra* (m) 'Schatz', cf. *vara* (m, n) 'Gegenstand des Wunsches'; *pārthyá* (m) 'Nachkomme des prthi'.

A pull to restructure *dyi-a* as *dy-ya* etc. would come from the well-known tendency in many languages for suffixes to grow at the expense of stems.

3.9. The curtailed treatment of *i*-stems is an innovation somewhere along the route between PIE and the language of RV. Vṛddhi is also an innovation somewhere along the same route—perhaps more or less simultaneous with the curtailed treatment. But the two developments appear to be independent, in that either would have taken place even without the accompaniment of the other.⁶

3.10. The theory of restructuring explains why *i*-stems alone, of all the PIE semivowels, came to receive the curtailed treatment. There was no chance that a derivative from a *u*-stem, say, would be restructured, even if the stem were to receive the shorter treatment; no chance, for instance, that **bhwa* 'tüchtig, kunstreich' from **bhú* 'kunstreich' would be restructured as **bh-va*; for there was no suffix *va* that was both productive and partially synonymous with *a*, and similarly for *ra*, *na*, etc. Nor was there any appreciable *u/a* alternation in existence. Thus there was no occasion for an **rbha* functioning as *a*-derivative of **rbhú* to arise.

4. DISCUSSION OF CERTAIN DETAILS.

To §3.3: *hṛdaya*. (a) The simplest PIE word for 'heart' is *kerd/gherd* (n). Whether these doublets are historically identical or are mere 'Reimwörter', Gdr.² 2.1.132, is immaterial so long as they behave precisely alike, as they seem to do. (b) There is also a well attested 'enlargement' of this, *kerdi/gherdi*, apparently used only in the nominative and the accusative. (c) It is a further fact that 'le nom de "cœur" est obtenu souvent au moyen de suffixes de dérivation comprenant *-i*...' (Ernout-Meillet). Many experts such as Ernout-Meillet, Wackernagel (3.34, 237, 305), and Renou (178, §229) cautiously content themselves with pointing out the coincidence between (b) and (c), but it seems legitimate to go further with Brugmann (2.1.157, 174, 199) and treat *hṛdaya* as reflecting PIE *ghrdejo*, an *a*- (PIE *e/o*-) derivative of *gherdi*. We would expect *ghrdejo*, so accented, to be primarily an adjective; the accent shift and the grammatical shift to a neuter substantive found in the historical RV word *hṛdaya* (n) are doubtless causally related to each other: cf. Burrow 27-8 and Renou 71 (lines 5-3 from bottom) on such accent shifts in general.

To §3.8: Concerning the *i/a* relationship in which the *a* comes from PIE *η*, there is a good deal to be said. Space and time prevent a thorough canvassing of the situation; I will illustrate it by the various related words for 'eye'.

In RV, *ákṣi* (n; also *akṣi*) and *akṣán* (n) occur as simplices, eight and ten times respectively, suppletively to each other in declension. Neither of them occurs as apparent final member of a compound. *akṣ* occurs only in *andakṣ* 'ohne Augen'; *akṣá* occurs only as the apparent final member of eight compounds. I say 'apparent' because the final *a* of e.g. *caturakṣá* should probably be regarded from a synchronic descriptive point of view as a *samāsānta* vowel (Wackernagel 2.1.101, §§44 ff.; Renou §163) which displaces *i*; in which case *caturakṣá* is not

⁶ See the brilliant articles of Kurylowicz cited in fn. 1, in which the structuralist's insight is applied, with success, to the old problems of *vṛddhi* and Brugmann's law. Incidentally, although several scattered cases of *vṛddhi* can be inferred for PIE, it can hardly have been a productive process there. Cf. Kurylowicz, *BSL* 44.43.

a compound but an *a*-derivative of an otherwise non-occurrent nonce-compound **caturakṣi*. Or if the *a* displaces *an*, then the nonce-compound is **caturakṣan*. In many instances the *saṁāsānta* element signalizes a derivative without any question; e.g. *ka*. So the use of *caturakṣā* in those cases in the declensional paradigm where *akṣi* would occur as a simplex, strengthens the conception of *-akṣā* as involving derivation from *akṣi*.

Further study is needed to determine how far *saṁāsānta a* was actually efficacious in encouraging the curtailed treatment of *i*-stems. There are complications, such as the fact that often *a* is the underlying stem vowel and *i* the *saṁāsānta* vowel, e.g. *paūrukutsi* 'Nachkomme des purukútsa' : *purukútsa*; *-gandhi* 'Geruch' : *gandhā*. And no rule is apparent as to when a compound will not and when it will be subjected to *saṁāsānta* derivation.

It is noteworthy that practically all sure instances of the curtailed treatment of *i*-stems in RV are of the *saṁāsānta* type. In other words, the statement of the grammarians that *i*-stems receive the curtailed treatment before secondary *a* is practically vacuous as far as RV is concerned. The sole contrary instance is *aurṇavābhā* 'von der Spinne stammend', and this is not quite perfect because *ūrṇavābhi* does not occur in RV. Then there are various instances that are not clear-cut; e.g. *nákta* (n) 'Nacht', alongside of *nákṭi* (f) 'Nacht' and *naktān* (n) 'Nacht'. Because of the accent and the meaning, there is no good reason either synchronically or historically for considering *nákta* to be an *a*-derivative of *nákṭi*; it is, however, possible to do so.

To §3.9: Latin *ovum*. Brugmann 2.1.158: 'Beachte das Fehlen des Formans *-i-* des zu grunde liegenden Nomens bei lat. *ovo-m*, zu *avi-s* ...' This implies both *vṛddhi* and curtailed treatment in Latin; it is not clear whether he means to ascribe it to PIE also: PIE *ōy-o* : PIE *ayi*. Other scholars, summarized in Walde-Hofmann 2.230 s.v. *ovum*, are equally evasive as to the PIE state of affairs. Although 'Zusammenhang mit *avis* ... ist wahrscheinlich' (Walde-Hofmann), it is not even sure which is the underlying word and which is derivative.

So it is far from clear that *ovum* implies curtailed treatment of *i*-stems in PIE; and even if it does, this case appears to be quite isolated, not an instance of a general practice.

5. RÉSUMÉ OF THE FACTS EXPLAINED. The preceding account is stronger at some points than at others. The facts which it is an attempt to explain should be borne in mind in any case:

(1) The curtailed treatment of *i*-stems is not PIE, apparently not even Proto-Indo-Iranian, but an Indic innovation.

(2) Secondary *ya*, in contrast to *va*, *ra*, *na*, etc., stands to secondary *a* in the unique relationship of being, at once, synonymous in large part and homonymous after *i*-stems (if differences in stem treatment be disregarded).

(3) The curtailed treatment of *i*-stems has few if any RV instances, apart from the *saṁāsānta* situations.

All of these considerations suggest that the curtailed treatment of *i*-stems is not only Indic, but is scarcely under way even in the Rigvedic period.

SOME HITTITE-SANSKRIT PARALLELS

E. ADELAIDE HAHN

Hunter College

I. HITTITE *-ma* AND SANSKRIT *smā*

There is a group of asseverative particles, usually postpositive, serving normally to stress the preceding word, which seem to be allied: Hittite *-ma*, Sanskrit *smā smā*, and Greek *μή*.¹

Hittite *-ma*, which is always enclitic, adds an adversative notion to the word which it follows, like Latin *tamen*, English *however*, *nevertheless*. Like all these words, it may also serve as a sentence connective, in which case its adversative

¹ Bibliographical references in this article are to be interpreted as follows. *Arch. or.* = *Archiv orientální*. Benfey = Theodor Benfey, *Handbuch der Sanskritsprache*, 2 vols.; Leipzig, 1852-4. Bennett = Charles E. Bennett, *Syntax of early Latin*, 2 vols.; Boston, 1910-4. Boisacq = Émile Boisacq, *Dictionnaire étymologique de la langue grecque*; Heidelberg and Paris, 1907. Brugmann, *Gdr.* = Karl Brugmann and Berthold Delbrück, *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*², 5 vols.; Strassburg, 1893-1916. Brugmann, *KVG* = Karl Brugmann, *Abrégé de grammaire comparée des langues indo-européennes*, transl. by J. Bloch, A. Cuny, and A. Ernout; Paris, 1905. Brugmann, *Totalität* = Karl Brugmann, *Die Ausdrücke für den Begriff der Totalität in den indogermanischen Sprachen*; Leipzig, 1893-4. Buck, *Comp. gr.* = Carl Darling Buck, *Comparative grammar of Greek and Latin*, 4th impr.; Chicago, 1948. Delbrück, *CO* = B. Delbrück, *Der Gebrauch des Conjunctivs und Optativs im Sanskrit und Griechischen*; Halle, 1871. Delbrück, *Gdr.*, see Brugmann, *Gdr.* Ernout and Thomas = Alfred Ernout and François Thomas, *Syntaxe latine*; Paris, 1951. Forrer, *BoTU* = Emil Forrer, *Die Boghazköi-Texte in Umschrift*, 2 vols.; Leipzig, 1922-6. Friedrich, *El.* = Johannes Friedrich, *Hethitisches Elementarbuch*, 2 parts; Heidelberg, 1940-6. Friedrich, *Vert.* = Johannes Friedrich, *Staatsverträge des Hatti-Reiches in hethitischer Sprache*, 2 parts; Leipzig, 1926-30. Goetze, *Tunnawi* = Albrecht Goetze, *The Hittite ritual of Tunnawi*; New Haven, 1938. Goodwin = William Watson Goodwin, *Syntax of the moods and tenses of the Greek verb*, rev. ed.; Boston, 1900. Güterbock, *Kumarbi* = Hans Gustav Güterbock, *Kumarbi*; Zurich and New York, 1946. Hahn, *Subj. and opt.* = E. Adelaide Hahn, *Subjunctive and optative: Their origin as futures*; New York, 1953. Hirt, *IG* = Hermann Hirt, *Indogermanische Grammatik*, 7 vols.; Heidelberg, 1921-37. Hofmann = J. B. Hofmann, *Etymologisches Wörterbuch des Griechischen*; Munich, 1950. Hofmann, *Lat. Gr.* = Manu Leumann and Joh. Bapt. Hofmann, *Stolz-Schmalz Lateinische Grammatik*⁶; Munich, 1928. Hrozný, *SH* = Friedrich Hrozný, *Die Sprache der Hethiter*; Leipzig, 1917. Ose = Fritz Ose, *Supinum und Infinitiv im Hethitischen*; Leipzig, 1944. Otten, *Mythen* = Heinrich Otten, *Mythen vom Gotte Kumarbi, Neue Fragmente*; Berlin, 1950. Pedersen, *Hitt.* = Holger Pedersen, *Hittitisch und die anderen indoeuropäischen Sprachen*; Copenhagen, 1938. Pedersen, *MS* = Albrecht Götze and Holger Pedersen, *Muršili's Sprachlähmung*; Copenhagen, 1934. Ribbeck = *Tragicorum romanorum fragmenta*³, ed. by Otto Ribbeck; Leipzig, 1897. Schwyzer, *Gr. Gr.* = Eduard Schwyzer, *Griechische Grammatik*, 2 vols. (2nd vol. compl. and ed. by Albert Debrunner); Munich, 1939-50. Sommer, *Bil.* = Ferdinand Sommer and Adam Falkenstein, *Die hethitisch-akkadische Bilingue des Hattušili I (Labarna II)*; Munich, 1938. Sturtevant, *Chr.* = Edgar H. Sturtevant and George Bechtel, *A Hittite chrestomathy*; Philadelphia, 1935. Sturtevant, *HG*¹ = Edgar H. Sturtevant, *A comparative grammar of the Hittite language*; Philadelphia, 1933. Sturtevant, *HG*² = Edgar H. Sturtevant (and E. Adelaide Hahn), *A comparative grammar of the Hittite language*, vol. 1,

force is often strong ('but, however'), sometimes weakened ('and, moreover'), as is the case also with Greek $\delta\acute{\epsilon}$ and Latin *autem*. Etymologically, *-ma* is probably connected with the conjunction *man* 'when, if' and with the modal particle *man -man*.²

The principal use of Greek $\mu\acute{\alpha}$ is as an asseverative particle in oaths, in combination with the name of a deity in the accusative case ('by'). In early Greek it was always postpositive, being preceded in affirmations by *val*, and in negations by *ov*. In one dialect (Thessalian) $\mu\acute{\alpha}$ also served, in place of $\delta\acute{\epsilon}$, as a mildly adversative sentence connective, always, like $\delta\acute{\epsilon}$, postpositive except in one late instance (2nd or 3rd century A.D.). Etymologically, $\mu\acute{\alpha}$ is probably connected with two other postpositive asseverative particles, $\mu\acute{\epsilon}\nu$ (usually but by no means always used in combination with its correlative $\delta\acute{\epsilon}$) and its stronger synonym, Homeric and Doric-Aeolic $\mu\acute{\alpha}\nu$, Attic $\mu\acute{\eta}\nu$.

It is obvious that semantically $\mu\acute{\alpha}$ as an asseverative particle is a close parallel for the Sanskrit enclitic asseverative particle *sma smā*, and as an adversative sentence connective for the Hittite *-ma*.

An etymological connection between Sanskrit *sma smā* and Greek $\mu\acute{\alpha}$ $\mu\acute{\alpha}\nu$ $\mu\acute{\epsilon}\nu$ is widely accepted, e.g. in the grammatical works of Delbrück (*Gdr.* 4.506-11), Brugmann (*KVG* 658, *Gdr.* 2.3.1008-9), and Schwyzler (*Gr. Gr.* 2.569), and in the etymological dictionaries of Boisacq, Hofmann, and Walde-Pokorny (2.685). Hrozný (*SH* 102 fn. 4) tentatively suggested a connection with $\mu\acute{\alpha}$ and $\mu\acute{\epsilon}\nu$ of Hittite *-ma*,³ and Sturtevant repeatedly equated Hittite *-ma* with Thessalian $\mu\acute{\alpha}$ (*Lg.* 6.30, *HG*¹ 87 and 135, *HG*² 44). He likewise in one instance equated Hittite *-ma* with Sanskrit *sma* (*HG*¹ 141)⁴ as an example of *s* movable; but this would seem to be ruled out by his later explanation (*Lg.* 14.241 fn. 11a) of Hittite *man* and *mahhan*, ingeniously identified with the *-o-* grade form of Greek $\mu\acute{\epsilon}\nu$ and with Greek $\mu\acute{\alpha}\nu$ respectively, as derived from an indefinite-relative stem beginning with *m-*, Indo-Hittite *me-*, Hittite *ma-*.⁵ Of course we cannot separate $\mu\acute{\alpha}$ from $\mu\acute{\epsilon}\nu$ and $\mu\acute{\alpha}\nu$; it must be from the Indo-European weak

rev. ed.; New Haven, 1951. Walde-Pokorny = Alois Walde, *Vergleichendes Wörterbuch der indogermanischen Sprachen*, rev. and ed. by Julius Pokorny, 3 vols.; Berlin and Leipzig, 1927-32. Whitney = William Dwight Whitney, *Sanskrit grammar*²; Cambridge, 1923. *ZA* = *Zeitschrift für Assyriologie und verwandte Gebiete*. References are to pages, unless there is indication to the contrary. Quotations from Hittite and Latin give only enough words to make clear the sense and the syntax, with no indication of omissions.

² A connection between the conjunction and the particle, though it is not universally accepted, seems to me exceedingly likely. See *Lg.* 18.103-4.

³ Also with Lycian *me*. Pedersen (*MS* 57-8) accepts this connection, but apparently thinks it rules out the possibility of a connection with $\mu\acute{\alpha}$; as for $\mu\acute{\epsilon}\nu$, this, he declares, 'muss wegen ganz abweichender Verwendung fern bleiben'. But $\mu\acute{\epsilon}\nu$ in the extended form $\mu\acute{\epsilon}\nu\tau\omicron\iota$ has the meaning 'nevertheless', and is also used (like Thessalian $\mu\acute{\alpha}$) as a correlative for $\mu\acute{\epsilon}\nu$ in place of $\delta\acute{\epsilon}$; thus it is unobjectionable as a semantic parallel for Hittite *-ma*.

⁴ Cf. also his earlier equation of them (*JAOS* 47.180-1); but here he was dealing not with the consonantism but with the vocalism, concerning which he expressed a view which he later retracted (*Lg.* 6.30 fn. 19).

⁵ The existence of an Indo-European pronominal stem in *m-* was demonstrated in 1938 by Sommer (*Bil.* 164-6) and Pedersen (*Hitt.* 67-71). Pedersen gives the stem as *mo-*, but presumably it had the two grades *mo-* and *me-*.

grade *m̃*. But then what of Sanskrit *sma smā*? I cannot believe that *μá* is to be separated from them either.

A solution is to be found if the explanation which I have made elsewhere (*Lg.* 18.83-116) can be accepted, namely, that the Hittite indefinite-relative-interrogative stem *ma-* is somehow connected with the familiar Indo-European indefinite stem *sem- som- sm- sm-*. In the article cited, I suggested (94-5) that we might recognize a case of *s* movable, as Sturtevant once did (*HG*¹ 141, already cited) concerning Hittite *-ma* and Sanskrit *sma*. However, in an excursus (113-6) I offered—tentatively and timorously to be sure, for, as I stated earlier (87), I realized ‘that such playing with building-blocks, while tempting and fascinating, is fruitless and at times perilous’—the alternative explanation that Brugmann’s proposal of years ago⁶ to see in *sem- som-* a combination of two elements, *se- so-* and *-mo-*, though never generally accepted, had at last been vindicated. Further vindication, for him and for me, is now provided by an article by van Windekens (*IF* 58.261-5) in which, mainly on the basis of Tocharian A *ša*, B *še* ‘one’, he expresses the belief, much as I had done on the basis in part of Hittite *šanaš* ‘one’ and its derivatives (*Lg.* 18.114-5),⁷ that Brugmann’s view⁸ is correct, and he concludes (265): ‘Es gab also im Indogermanischen eine einfache Wurzel **se-* “eins, zusammen”: durch Erweiterungen wurde diese zu **se-m-*, das in fast allen idg. Sprachen die Oberhand gewonnen hat und die das idg. **se-* vertretenden Formen verdrängt hat’.⁹ That he and I independently,¹⁰ on the basis of two quite remote languages, reached the same conclusion, seems to me to provide confirmation for both of us.¹¹ Then Hittite *-ma* belongs with the simple stem in *m-*, and Sanskrit *sma* with the composite stem in *sm-*; Greek *μá* is of course ambiguous, but probably, if the Greek forms are as close parallels to the Hittite as Sturtevant holds, goes with the *m-* group.

A word needs to be said here about the meaning of Hittite *šanaš*. Goetze has so completely convinced me of the correctness of his demonstration (*Lg.* 11.185-90, *Tunnawi* 127) that *šanaš* = ‘one’, that I refuse to be convinced by his later retraction (*Arch. Or.* 17.296): ‘The reexamination of *šan(n)i/a-* and its derivatives, then, leads to the conclusion that its basic meaning is “single, one and the same”. ... The present demonstration implies a slight modification of my former

⁶ He had originally suggested something of the sort in *Totalität* 49 fn. 1 (1893-4); later apparently retracted this at least so far as *ἐκάρων* went, *IF* 21.8 (1907); but finally returned to it, *IF* 37.160-1 (1916-7). Brugmann of course did not recognize the element *-mo-* as an indefinite stem; he identified it with the superlative ending (ib. 159). But, as I noted in the article cited (114), he was making a comparison more apt than he realized when he linked Sanskrit *sama-*, Greek *ὁμός*, and Latin *similis* with Greek *αἰρός* and *αἰρός*, Latin *aequus*, in which he identified the first element with a demonstrative and the second with the indefinite stem *kwo-* (156-9).

⁷ We also agree in suggesting a possible connection with the reflexive pronoun (*IF* 58.265 fn. 1, *Lg.* 18.115-6).

⁸ As expressed in *Totalität*, cited 264 fn. 4. He does not mention the article in *IF* 37.

⁹ This of course accounts neatly for the much disputed forms Greek *ἐκάρων* (cf. sup., fn. 6) and Sanskrit *sahasra-* (265).

¹⁰ Both our papers appeared in 1942.

¹¹ Schwyzler (*Gr. Gr.* 2.569 fn. 3) seems willing to accept our view.

view. In particular I concede to my opponents that *šanaš is not the reading of the numeral 1-aš.'

Since both 1-aš and 1-iš exist, the word for 'one', whatever it is, must exist in two (probably related) forms, one ending in -aš and the other in -iš. Goetze himself mentions the possibility (ib. fn. 43) that the form in -iš may be šannapiliš or *šanaiš. Of šannapiliš he has already said (296) that it 'seems to have the tendency to supersede the more archaic šani/a-'; but I think šannapiliš usually has a more emphatic and more specific meaning than 'one', corresponding to šanaš as Latin *sōlus* or *singulus* does to *ūnus*.¹² On the other hand šanaiš could easily be a by-form for šanaš.¹³ I once suggested (Lg. 12.115) that tamaiš 'other', the second element in which I proposed to equate with the Indo-European suffix -mo- seen in ordinal numerals and superlatives, might have been formed from an original *tamaš under the influence of hantezziš 'first' (which itself alternates with a variant in -aš, hantezziyaš); and exactly the same explanation could apply to *šanaiš beside šanaš.

However, though Goetze is willing to consider the possibility that 1-iš is *šanaiš, he dismisses the possibility that 1-aš is šanaš, because of the support lent by the Kumarbi text to the supposition 'that the numeral for "one" should be read ašma- in Hittite' (297). I question this. Goetze (296) quotes the sequence a-aš-ma-at-ta — da-an-ma-at-ta — 3-an-na-at-ta. This occurs in KUB 33.120.1.31-3 (Güterbock, *Kumarbi* *2). We also meet a-aš-ma-an in KUB 33.106.4.13 (*Kumarbi* *28). Güterbock refers to the form as ašma- in his note on the first passage (p. 35), but demonstrates that it is ašma in his note on the second (p. 79),¹⁴ and this is confirmed by the occurrence of a-aš-ma-wa in a more recently discovered fragment (see Otten, *Mythen* 10 fn. 2). Thus with all enclitics removed, the series of numerals appears as a-aš-ma da-an 3-an-na 'first, second, third'.

I agree with Otten, KUB 33 p. iv, that all are ordinals. Goetze (*Arch. Or.* 17.297 fn. 46) takes da-an and 3-an as ordinals, but (citing Akkadian usage) suggests that ašma-, as he calls it, is the cardinal, hence 'one', because the ordinal is hantezzi-. However, hantezziš is the ordinal ADJECTIVE; ašma, like dan and 3-anna, is an ADVERB, 'first' in the sense of the sometimes used 'firstly'. On the other hand, when we say 'for the first time', we need an adjective for 'first', ha-an-te-iz-zi BAL-ši (found e.g. KBo 3.2.1.3 and 40, KUB 19.37.3.22; cf. Hatt. 2.30). Thus Forrer is surely right in his suppletion in KBo 3.16.2.1 (= 2BoTU 4a.2.1), so that we read (ib. 1-3) [ha-an-te-iz-]zi BAL-ši, ta-a-an, 3-na. This seems to prove that in a series for 'first, second, third', Hittite, unlike Akkadian, uses an ordinal for 'first' as well as for 'second' and 'third'; hence we may conclude that ašma is an ordinal adverb meaning 'first(ly)',¹⁵ and we need not abandon the

¹² So far as derivation goes, I have suggested (Lg. 18.91) a parallelism with Latin *simplex*.

¹³ On the commonness of -is adjectives in Hittite, cf. Sturtevant, Lg. 10.266-73.

¹⁴ Approved by Friedrich, ZA 15.244.

¹⁵ The etymology of ašma is problematical. One thinks of the possibly similar formation -ašta 'then, finally' (cf. našta = nu + -ašta), but this is always enclitic. Can the -ma part be the element which, as I have noted above, I have already suggested (Lg. 12.115) is perhaps to be recognized in tamaiš 'other', and which, incidentally, is the one that Brugmann

view so ably set forth by Goetze that the cardinal adjective meaning 'one' is *šanaš*.

II. DOUBLE DATIVE: INFINITIVE PLUS DATIVE NOUN

The Hittite verbal noun, regularly a neuter *-r/n-* stem, is much more a noun than a verb. As a noun, it has case; but as part of a verb, it lacks most of the characteristics of finite forms,¹⁶ being voiceless and tenseless, never having a subject and very rarely an object.¹⁷ Verbal nouns in the dative case regularly, as befits a dative, function as expressions of purpose. There are two types of these, absolutely parallel in use, those in *-anna* or *-anni*, from nominatives in *-tar*,¹⁸ and those in *-manzi* or *-wanzi*, somehow allied to nominatives in *-mar/-war*. It is convenient to call both types infinitives;¹⁹ there seems no more reason to class them in different categories than there would be to separate the Latin passive infinitive in *-ī* from those in *-ārī*, *-ērī*, and *-īrī*.

The primarily nominal character of the Hittite infinitive virtually compels certain syntactic phenomena. If we wish to say in Hittite 'he collects the troops', we use the same construction as in English; the substantive corresponding to 'he' (if used) is the subject, and that corresponding to 'troops' the object, of the verb corresponding to 'collects', precisely as in Latin, (*is*) *copias cogit*. But if we wish to say 'he directed him to collect the troops', we cannot say anything comparable to the Latin *iussit eum copias cogere*, in which *eum* serves as subject, and *copias* as object, of the infinitive *cogere*;²⁰ the words for 'him' and for 'troops'

posited (as noted above, fn. 6) as the second element in the root *sem-* (IF 37.159) and that I have proposed to equate with the indefinite stem (Lg. 18.114)? This element occurs frequently in Indo-European in the formation of both ordinals and superlatives; cf. e.g. the word for 'first' in many languages, as Sanskrit *prathama-* and Latin *prīmus*. That it functioned in Hittite as an ordinal suffix is proved by the dative *ši-ip-ta-mi-ya* 'seventh'.

¹⁶ As do also the verbal adjectives (participles).

¹⁷ As instances of this very exceptional usage, we may cite KBo 4.4.2.63-4 *ku-e KARAŠ.HI.A I-NA KUR^{URU}Nu-haš-ši hal-ki.HI.A-uš har-ni-in-ku-wa-an-zi pī-e-hu-da-an har-ta* 'the troops which he had led into Nuhasi in order to destroy the crops', and KUB 24.5.2.17-8 IGI.HI.A-YA-wa-za-kán tu-uk ú-wa-an-na ha-aš-ši-ik-ki 'open my eyes for seeing you', i.e. 'open my eyes so that I may see you'. A few other examples of the combination of case-forms of nouns with an infinitive are given by Ose, 57 and 86.

¹⁸ These are allied to the Latin gerund. See Sturtevant, Lg. 20.206-11, and HG² 73.

¹⁹ It has been customary to call the forms in *-manzi/-wanzi* supines; so e.g. Friedrich, EL. 1.80. But to me it seems preferable to bestow the name of infinitive on both types; so e.g. Sturtevant, HG¹ 267 and HG² 148. Cf. Ose 88.

²⁰ The Latin construction must have been in origin precisely parallel to the Hittite: thus such a passage as Terence, Heaut. 585-6 *iube hunc abire* must have meant originally 'drive him for the purpose of going', with *hunc* as object of the main verb and *abire* a dative of purpose. But the Latin infinitive eventually lost its dative force, becoming purely an indeclinable verbal noun which functioned mainly as nominative and accusative (with the inflected gerund taking over the other case relations); and in this new development it acquired the ability to possess not merely an object (*copias*) as the Hittite infinitive occasionally did (cf. fn. 17), but even a subject, the accusative which had once served as object of the main verb. But this state of affairs could apply only to the infinitive in combination with a main verb of ordering, compelling, or allowing; the use of the infinitive with subject-accusative in combination with a main verb of saying or thinking or perceiving, in connection with which the infinitive could never have had dative force, must have had a quite different origin (see my discussion of this in TAPA 81.117-29).

cannot be separated from the main verb,²¹ and the result is something which can be rendered literally into English only by some such unidiomatic phraseology as 'he directed him to the troops for collecting': Hatt. 4.4 *na-an-kán A-NA ERÍN.MEŠ ni-ni-in-ku-u-an-zi ú-e-ri-ya-at*. Both *A-NA ERÍN.MEŠ* (the composite Sumerian-Akkadian writing of Hittite *tu-uz-zi-ya*) and *ni-ni-in-ku-u-an-zi* are datives.

The parallelism of the Hittite construction with that of Vedic at once becomes obvious. The combination of a dative noun of reference and a dative infinitive of purpose occurs in Vedic just as in Hittite. As examples²² may be cited RV 5.2.9 *śiṣṭe śrṅge rakṣase vinikṣe* 'he whets his horns for the demon for piercing', i.e. 'he whets his horns to pierce the demon', RV 10.14.12 *asmabhyam dṛśaye sūryāya punar dātām asum* 'let them give us life again for the sun for seeing', i.e. 'let them give us life again for seeing (that we may see) the sun'.²³ The Vedic construction like the Hittite must present the original state of affairs. Thus I believe it is methodologically indefensible to describe it as if it were an anomalous and later-developed substitute, perhaps due to attraction and certainly illogical, for the normal construction (i.e. normal in English or German!). As examples of this sort of approach I cite the following descriptions. Brugmann, *Gdr.* 2.3.917: 'Abhängigmachung des dem Infinitiv untergeordneten obliquen Substantivkasus vom regierenden Verbum. Es handelt sich hier um eine Art Antizipation.' Delbrück, *Gdr.* 4.470: 'Endlich kann noch neben dem zur Ergänzung der Satzaussage dienenden Infinitiv der als Objekt empfundene Substantivbegriff im Dativ erscheinen.' Whitney 352: 'The noun which is logically the subject or the object of the action expressed by the infinitive is frequently put beside it in the dative (by a construction which is in part a perfectly simple one, but which is stretched beyond its natural boundaries by a kind of attraction)'. Kirk, *TAPA* 73.296: 'In Vedic Sanscrit a substantive which is properly the object of a dependent infinitive may be shifted to dependence on the main verb, and the Vedic infinitive, having different case forms, may be assimilated in case to its former object.'

Benfey (1.432) compared the Vedic usage with the Latin gerundive construction (or future passive participle as he called it). Delbrück at first objected to this view (*KZ* 18.104) because he believed that the two datives are independently used; but later, presumably because of his realization that the dative noun is found only in combination with a dative infinitive (*Gdr.* 4.470-1), he withdrew his objections and accepted the identification of the Vedic infinitive with the Latin gerundive (ib. 471), as did also Brugmann (*Gdr.* 2.3.918, 922). But the comparison with the gerundive seems to me completely incorrect: the Vedic infinitive is a singular neuter noun, and agrees with its accompanying substantive in case only; the Latin gerundive is an adjective, and of course agrees with

²¹ When the main verb is in the passive, the word which would be the object of the infinitive in English is attached to the main verb as its subject. Thus if we wish to say 'it was directed to sacrifice a bull to him', we must recast this as 'a bull was directed to him for sacrificing', *MS* 1.11-2 *nu-uš-ši GUD pu-u-hu-ga-ri-iš pt-i-ya-u-wa-an-zi SI × DI-at*.

²² For other examples, see Brugmann, *Gdr.* 2.3.918 and Whitney 352.

²³ Both Brugmann (*Gdr.* 2.3.918) and Delbrück (*KZ* 18.104 and *Gdr.* 4.470) seem to think the position of the dative of reference—whether before the infinitive as in 5.2.9, or after it as in 10.14.12—is of some significance; but this I cannot see.

the substantive that it modifies in gender and number as well as in case. Brugmann's comparison (*Gdr.* 2.3.918) of the Vedic construction with the Latin *decemviros legibus scribendis creavimus* is utterly wrong; the only true parallel would be **decemviros legibus scribendo creavimus*. For this I can offer no indubitable justification, but I am sure it represents the original structure of such a locution as we meet e.g. in Plautus, *As.* 250 *argento comparando fingere fallaciam*. Here I believe *comparando* was originally a noun quite as much as *argento*, the meaning being 'to devise a trick for money, for obtaining', i.e. 'for obtaining money';²⁴ since Latin doubtless from an early date possessed adjectives in *-ndus*, such as *oriundus*, *volvendus*, *secundus*, and *rotundus*, most of them associated with verbs, it was easy to assume that *comparando* was also an adjective like them, in agreement with *argento*.²⁵ Thus from the gerund construction developed the gerundive construction.²⁶ This shift would of course have been possible only at an early date when verbal nouns and adjectives were still, as always in Hittite, voiceless. That the phenomenon was not merely Latin but Italic is proved by the existence of the gerundive in Oscan and Umbrian; the absence of the gerund in the relatively scanty remains of these languages can easily be due to chance.²⁷

The examples in Hittite, Vedic, and Latin just discussed all constitute one particular manifestation of the familiar 'double dative' construction, consisting of a combination of a dative of reference (generally denoting a person) and a dative of purpose (often an abstract or verbal noun), so common in Latin,²⁸ e.g.

²⁴ So, rightly, Hofmann (*Lat. Gr.* 597), who translates 'für das Geld, seine Beschaffung'.

²⁵ Once the shift was made, the gerundive might appear in passages where the gerund would scarcely have applied, as in Plautus, *Pers.* 5 *fio miser quaerendo argento*.

²⁶ I offer this theory and discuss it at length in *TAPA* 74.277-98.

²⁷ Of course the foregoing hypothesis would be proved only if we actually possessed in Latin such a passage as **decemviros legibus scribendo creavimus*, but we do not; where the noun is feminine or plural, we find no instance of a gerund, only the agreeing gerundive, as in Brugmann's *decemviros legibus scribendis* or in *triumviri rei publicae constituendae*. But in the genitive there are plenty of instances of the gerund: e.g., Plautus, *Capt.* 1008 *lucis tuendi copiam*, ib. 852 *nominandi istorum copia*. Sporadic vestiges are met in classical times, especially with the plural *sui*, perhaps because it looked like a singular, as in Caesar, *Bell. Gall.* 7.43.2 *legatos sui purgandi gratia mittunt*, and with the genitive plural in *-ōrum* or *-ārum*, probably to avoid cacophonous repetition, as in Cicero, *Inv.* 2.2.5 *exemplorum eligendi potestas*. But infinitely commoner was the shift to the gerundive construction, from *lucis tuendi* to *lucis tuendae*, undoubtedly, in my opinion, brought about by the ambiguous form in the masculine or neuter, as Accius 275-6 (Ribbeck) *luminis conspiciendi insolentia*.

²⁸ But there are two important differences between the gerund-gerundive construction just discussed, and other manifestations of the double dative in Latin. (1) The double dative in Latin (and indeed the use also of a dative of purpose alone without a second dative) seems more restricted than in Hittite or Vedic. It is mainly confined (a) to verbs of being, and (b) to verbs with the general notion of giving, assigning, or deeming, either active or passive. It is closely related either to the subject (with the copula and with passive verbs) or to the object (with active transitive verbs); indeed it may interchange with the predicate nominative in the former case and with the predicate accusative in the latter. (2) With a gerund the performer of the action is regularly either the subject of the verb modified by the gerund, or a generalized impersonal agent; and the accompanying dative substantive represents the recipient of the action. Similarly with the ordinary double dative construction, where we have a noun of action corresponding to a transitive verb, the dative substantive usually designates the recipient of the action, as in Plautus, *Mōst.* 922 *ne quid captioni mihi sit*; but there are exceptions in which it designates the

Terence, Eun. 135 *emit eam dono mihi* 'he bought her as a present for me' and the often-quoted legal and proverbial phrase *cui bono?* This double dative in its turn is probably one manifestation of what may have been a much more widespread and basic phenomenon—the so-called partitive apposition or *σχῆμα καθ' ὅλον καὶ μέρος*. Latin provides us with plenty of double datives of this category.²⁹ The most clear-cut instances are those involving parts of the body, as in Plautus, Cas. 337 *quis mihi subveniet tergo aut capiti aut cruribus?* and Curc. 486 *linguae moderandum est mihi*;³⁰ then, on a slightly less concrete plane, the *animus*, which can have (1) physical connotations, as in Mil. 1331–2 *animo male factum est huic repente miserae* and Amph. 1057–8 *vae miserae mihi animo malest* (cf. the amusing transference to a building, Pseud. 952–3 *animo malest aedibus* 'the house is sick'), and (2) mental connotations, as in Stich. 524 *si tibi nullast aegritudo animo obriam*. From this second use of *animus*, it is but a step to the purely abstract *ingenium*, in Capt. 371 *tute tibi tuopte ingenio prodes plurimum*; thence to other abstract nouns, as in Trin. 313 *istaec ego mi semper habui aetati*³¹ *integumentum meae* and Bacch. 1083 *nimi' nolo desidiaei dare ludum*; and finally to an abstract noun of action, as Bacch. 439 *magistro desinebat esse dicto oboediens*.

There is reason to believe that partitive apposition was much commoner at an early stage of the language than later. It abounds in Hittite.³² As examples with the dative may be cited Al. 4.45 (Friedrich, *Vert.* 2.82) *nu-kán A-NA* ^{PUTU-ŠI} *ŠU-i an-da a-aš-šu lu-ú-lu a-ú* 'and in my Majesty's hand (i.e. protection) behold goodly prosperity',³³ Instr. 1.20 (Sturtevant, *Chr.* 148) *nam-ma-kán pt-di KÁ-aš li-e ti-ya-zi* 'let him not approach the place, the doors', i.e. 'the doors of the place'. Other cases are commoner than the dative.

agent instead, as in Trin. 632 *ut sic odio esses mihi*. This exception seems to be the rule with one special type of noun of action, the verbal noun (really a dative supine) in *-tui* or *-sui*, as in Curc. 501 *bono usui estis nulli*.

²⁹ Of course partitive apposition in Latin is by no means confined to the dative. See Hofmann's extremely interesting discussion and collection of examples, *Syntaktische Gliederungsverschiebungen im Lateinischen infolge Erstarrung ursprünglich appositioneller Verhältnisse* (*IF* 42.81–5). Hofmann cites numerous examples not only of the dative (84–5) but also of the nominative (85–6) and of the accusative (81–4), a few of the ablative (85) and of prepositional phrases (86), but none of the genitive; however, he suggests (85 fn. 2) that the construction *lucis tuendi* (discussed above, fn. 27) may exemplify this type. I formerly doubted this (see *TAPA* 74.282); but I now realize, as I have just said, that the fundamental relationship of the verbal noun to the substantive which it accompanies may be viewed as in origin one of quasi-apposition.

³⁰ Some might term *mihi* a dative of agent, but actually the so-called dative of agent is essentially a dative of reference.

³¹ On the use of this particular noun in partitive apposition in the dative, cf. its similar use in the accusative, illustrated below, fn. 34.

³² For discussion and examples, see Friedrich, *ZA* 1.175 and 2.47, *Vert.* 1.43–5 and 2.24, *El.* 1.69; Sommer, *Bil.* 106.

³³ It is hardly possible to render this literally. The force of the dative *A-NA* ^{PUTU-ŠI} must be 'with reference to my Majesty' and that of *ŠU-i* 'with reference to (my Majesty's) hand', but the adverb or postposition *anda* adds a locative notion 'in, within' at least to *ŠU-i* and perhaps to *A-NA* ^{PUTU-ŠI} also. (It must not be forgotten that the Hittite dative case serves as a locative as well.)

Many constructions owe their genesis to early instances of partitive apposition. For instance, the accusative so used³⁴ may have given rise to the so-called accusative of specification and accusative used as object of the middle voice³⁵ met both in Greek authors and in Latin poets (like Vergil) who imitate Greek. Of course such an accusative must have started as the object of an active verb, but by a misunderstanding of the original construction remained as an accusative with a passive verb or an adjective, the transition possibly being achieved through the use of a middle verb, since this might be equivalent to an active verb with an accusative reflexive.

The scheme of part and whole as we know it usually involves concrete nouns, generally denoting, as we have already mentioned, a person and a part of his body. But we have seen in our Latin examples the possible introduction, as the member representing the part, of an abstract noun, one signifying a quality or even, as in Bacch. 439, an action. If a verbal noun of the type of *dicto* can be placed in apposition with the substantive denoting the person uttering the *dictum*, why not, at an early stage, any verbal noun whatsoever with the substantive denoting the person or thing in some way connected with the action denoted by the verbal noun, whether as performing the action involved ('subject') or as receiving it ('object')?

This possibility throws fresh light on all the Vedic infinitives and all the Latin gerunds. It is now easy to see why the construction that we have been studying is not confined to the double dative type. Just as Latin possesses a genitive and an ablative use of the gerund-gerundive construction, so Vedic possesses a combination of the genitive infinitive beside a genitive substantive and of the ablative infinitive beside an ablative substantive, parallel to the double dative.³⁶ It is possible that the double accusative also enters into the picture. Vedic infinitives differ from Hittite infinitives in two important respects. (1) They include accusative forms³⁷ as well as datives;³⁸ these have two main uses, as the object of an auxiliary verb, like the Latin infinitive after it lost its dative force, and in expressions of purpose with a verb of motion, like the Latin ac-

³⁴ As in Homer, Il. 7.14-6 *Ἰφίτοον βάλε δουρὶ δουρὶ* and Plautus, Rud. 1345-6 *te Venus eradicet caput atque aetatem tuam*. With the latter instance of apposition of part and whole, contrast their coordination in the same play 486 *Neptuno credat sese atque aetatem suam*; and for the appositional use of the accusative *aetatem*, cf. the similar use of the dative *aetati* in Trin. 313, already cited above (see fn. 31).

³⁵ The constructions to which these names are given seem to me to be essentially the same, except that with a verb or a participle (e.g. Vergil, Aen. 1.561 *voltum demissa*) either name can be applied, and with an adjective (e.g. ib. 320 *nuda genu*) only that of an accusative of specification. It is significant that both usually indicate a part of the body, or, by an extension (as in Vergil, Ecl. 1.53-4 *saepes apibus florem depasta*), some portion of a whole.

³⁶ For examples, see *Gdr.* 2.3.919. Brugmann adds to his Vedic instances parallel examples from Avestan, and then goes on to cite examples from Greek (920-1) and from Balto-Slavic (921-2).

³⁷ Rather oddly, these are the only ones that survived into the classical period.

³⁸ In Hittite, there is a supine in *-wan* which is regularly classed as a suffixless dative (or locative), but which I have sometimes been tempted to regard as an accusative; it is restricted to a very limited use in combination with not more than two or three auxiliary verbs, in the sense of 'begin'.

cusative supine in *-tum*, with which the Sanskrit accusative infinitive is cognate. (2) They take an accusative object with far more freedom than the Hittite infinitive. Is there perhaps a connection between these two phenomena? Is it possible that the use of the accusative as the object of the infinitive had its genesis in an accusative substantive used in partitive apposition with an accusative infinitive, whence it spread to infinitives in other cases? Thus *pāsān vicrtam vettha sarvān*, translated by Whitney (352) 'thou knowest how to loosen all bonds', would have originally meant 'thou knowest all bonds, (their) loosening', etc.

There are still further implications. Since the verbal noun, being originally voiceless, may signify either an action performed by the accompanying substantive or one received by it, in other words since the substantive may have the relation of either subject or object to the verbal noun, we have the possibility that the double accusative construction may have a bearing on the development in Greek and Latin of the infinitive with subject-accusative.³⁹ This is in line with the view of Ernout and Thomas, who believe that both the substantive that later became the subject of the infinitive, and the infinitive itself, were originally objects of the main verb. Thus they analyze *sentio eum venire* as a combination of *sentio eum* and *sentio venire* (271-2; cf. 274). However, on the whole this hypothesis seems to me dubious; would we not expect *sentio veniendum* (gerund)⁴⁰ rather than *sentio venire*? It must not be forgotten that the Latin (and Greek) infinitives were originally datives, not accusatives; they show their dative force clearly in the construction *iubeo (sino, cogo) eum venire*, and it does not seem likely that they would have lost it so completely at so early a period in the construction *scio (credo, dico) eum venire*. Furthermore, though *eum* can be regarded as a possible object of *scio*,⁴¹ it seems hard to envisage it as an object of *credo*, and still more so as an object of *dico*; *dico eum* simply does not make sense.

III. INDICATIVE WITH A MODAL PARTICLE IN PROHIBITIONS

Indo-Europeanists, accustomed to the use in Greek and Latin of the subjunctive in a prohibition, have been puzzled by the corresponding use of an indicative (an augmentless past form) in Sanskrit. To explain this use, they invented a special mood called by Delbrück (*CO* 5) the 'improper subjunctive' ('*unechter Conjunctiv*') and by Brugmann (*MU* 3.2 and fn. 3) the 'injunctive' ('*Injunktiv*'); and almost without exception they posit its existence in Indo-European as a parallel to indicative, subjunctive, optative, and imperative. The only exceptions known to me are Hirt, who urges (*IG* 6.267) that the injunctive be explained simply as an indicative with subjunctive sense, and Buck, who (*Comp. gr.* 238) warns against 'supposing that this is a distinct formal category, coordinate with

³⁹ In that case the predicative construction that I view as the starting-point of the locution would have been merely a contributing, not a determining, factor in its development. Cf. sup., fn. 20.

⁴⁰ In classical Latin, to be sure, the gerund is limited to the use with a preposition, but in early Latin it is freely employed as a nominative-accusative; see *TAPA* 74.286-94.

⁴¹ Cf. the suggestion made at the close of the preceding paragraph about the possible relationship of *pāsān* to *vettha*.

the other moods'. Indeed, how could it be viewed as 'a distinct formal category', since in form it is purely, as Hirt puts it (*IG* 6.252), 'ein ganz gewöhnlicher Indikativ'? The absence of an augment certainly does not distinguish it from other indicatives; the use of the augment seems to be limited to three language-groups, Indo-Iranian, Armenian, and Greek, and in all these its practically universal use became obligatory only in later forms of the languages, and so must have been an independent development in each case. Whitney's discussion of the phenomenon seems to me to show less than his usual acumen; he says (211): 'the name of subjunctive, in the forms "imperfect subjunctive" and "improper subjunctive", has been also given to the indicative forms of imperfect and aorist when used, with the augment omitted, in a modal sense ...: such use being quite common in RV., but rapidly dying out, so that in the Brāhmaṇa language and later it is hardly met with except after *mā* prohibitive.' Actually, it is this sole surviving use that must have been the original one, and it is with this that our investigation must start.

Those who have dealt with the Sanskrit form of prohibition have been, in my opinion, led astray by their assumption that what may be termed modal relations are of necessity expressed by special moods of verbs. This has ceased to be the case in some modern languages; thus English uses the so-called modal auxiliaries rather than moods. And I believe it was not the case in the original language either, though here recourse was had to a quite different type of expression, namely, modal particles. Hittite had only two moods: the indicative, used in factual expressions, and the imperative, used in commands; other shades of meaning were conveyed by the use of modal particles in combination with an indicative, *man* in expressions of potentiality or unreality, *le* in prohibitions. This *le* corresponded to the ordinary negative particle, *natta*, as did Sanskrit *mā*, Greek *μή*, and Latin *nē* to *na*, *οὐ*, and *nōn* respectively. The use of one of these negative modal particles in conjunction with a special verb mood-form, subjunctive or optative, involved a tautological innovation which doubtless arose by analogy as the subjunctive or optative came to be substituted for an indicative in affirmative expressions. This tautology on the part of the ancient languages has persisted in modern Greek, although here the confusion of *οὐδέ* and *οὔτε* with *μηδέ* and *μήτε* shows that the original distinction is no longer clear cut; in the Romance languages the situation has been conveniently and sensibly simplified by the disappearance of the dichotomy in the use of negative expressions and the survival of only a single negative particle.

Thus in the study of prohibitions it is the use in Greek or in Latin of the modal particle *μή* or *nē* with a subjunctive that needs a special explanation, and not the use in Vedic and classical Sanskrit of the modal particle *mā* with the indicative.⁴² Equally easy to understand is the use of the non-modal particle *na* with the

⁴² In early Latin we find a possible parallelism in the use of certain verb forms which, like the Greek second aorist, combine a weak or aorist stem with the endings of the imperfect—these endings in the Latin forms being *-am -ās* etc., seen not only in the imperfect but also in the pluperfect. These are surely past indicatives in origin. Instances are *ne fuas* (Plautus, *Capt.* 443 and *Trin.* 287) and above all the very common *ne attigas, ne attigatis* (Plautus, *Bacch.* 445, *Ep.* 723, *Most.* 468, *Truc.* 276; Terence, *And.* 789; *CIL* 1².424 Numbers 499 and 500).

optative which appears in the Veda, and later becomes common;⁴³ this is presumably an independent development in Sanskrit, resulting from the acquisition of modal force by the optative mood, which thus rendered the modal particle otiose and unnecessary and led in this particular construction to its disappearance, comparable to its disappearance, already referred to, in Romance centuries later. The extremely rare use in Sanskrit of the optative with *mā* must be due to a tautological contamination of the sort that became the norm in Greek and Latin.⁴⁴ (A type of contamination of a different sort is exemplified by the use of the imperative with *mā* which begins to appear in the epics; similarly we have a few cases of the imperative with *le* in Hittite,⁴⁵ a number with $\mu\eta$ in Greek, and some with *nē* in Latin.⁴⁶)

The use in prohibitions of an augmentless past indicative has thus been accounted for, so far as the mood goes. But why past and why augmentless?

The use of the past tense must in my opinion go back to a period when the 'tenses' expressed aspect rather than time. There seems to be a widespread tendency in both commands and prohibitions to emphasize the punctual rather than the iterative or durative phase of an action: a person is directed to do something once and for all, or not to do it even a single time, rather than to do it, or not to do it, repeatedly or continuously.⁴⁷ In Hittite, though in prohibi-

⁴³ Exactly parallel is the use in prohibitions in early Latin with the subjunctive of *nōn* or *neque* instead of *nē* or *nēve*. We find *neque* fairly often, but *nōn* is extremely rare; Bennett (1.170) cites only two instances.

⁴⁴ The development of modal relations in affirmative expressions is similar. Hittite uses *man* with a so-called present (really present-future) indicative to express potentiality, and with a past indicative to express unreality. In Homeric Greek, which was certainly a transitional period in which modality was coming to be expressed by verbs rather than by particles, we find considerable fluctuations of usage; potentiality or remote futurity can be expressed by the indicative (future) with a modal particle ($\alpha\epsilon$ or $\alpha\upsilon$) or by a mood (subjunctive or optative, though the subjunctive denotes more vivid rather than less vivid futurity) without a particle or, tautologically, with it; unreality can be expressed by the past indicative with a particle or, tautologically, by the optative with a particle. Eventually in Greek the use of the past indicative (plus a particle) became the sole norm in expressions of unreality, exactly as in Hittite; but in expressions of potentiality, the standard form was the tautological one, optative plus particle. Sanskrit and Latin, on the other hand, use no corresponding affirmative modal particle, and depend purely on moods for the expression of modal relations so far as potentiality and unreality go.

⁴⁵ Cf. Sturtevant, *HG*¹ 251 fn. 70.

⁴⁶ The Greek and Latin imperatives are used only in a restricted sense; see fn. 47.

⁴⁷ When the iterative or durative aspect of an action needs to be stressed, for some reason both Greek and Latin show a tendency to use the imperative (with $\mu\eta$ and *nē* respectively). In Greek the regular use in prohibitions is the present imperative is met very seldom, and the aorist subjunctive for punctual aspect; the aorist imperative is met very seldom, and the present subjunctive not at all (see Goodwin 89). In Latin the perfect subjunctive is regularly punctual and the present subjunctive is often, though not always, durative, as in the common phrases *molestus ne sis* (Plautus, *As.* 469 et al.) and *ne vereare* (Capt. 349 et al.); Delbrück (*Gdr.* 4.383) thinks this distinction is the rule, and, though Bennett is probably right in considering this generalization too sweeping for early Latin so far as we know it (1.174), it doubtless is correct, as Bennett himself agrees (ib.), as applied to the original distinction between the two tenses. As for the present imperative in prohibitions (as *ne fle*, *ne time*), this certainly does seem, precisely like the present imperative in Greek, to be confined to durative use; cf. Ernout and Thomas (197), 'il

tions it is the present indicative that is combined with *le*, in orders the form used, the imperative, has secondary endings, as have also a number of imperative forms in Sanskrit, Greek, and Latin. In prohibitions in Greek, we find only the aorist subjunctive, not the present, with $\mu\acute{\eta}$.⁴⁸ In prohibitions in Latin, the perfect subjunctive as well as the present is met with *nē*,⁴⁹ and there are even a few possible examples of it in commands.

When the shift in the force of most tenses from aspect to time took place, the one form of locution completely and universally unaffected by the change was of course that used in commands and in prohibitions. One cannot issue orders, whether positive or negative, with respect to the past. Hence it was natural that not only in Greek, where everywhere outside the indicative the difference between aorist and present remained a matter of aspect, but in Sanskrit and Latin as well, secondary forms remained in imperative and prohibitive expressions. Naturally the augment that was used in Sanskrit and Greek to mark past time would not apply here; hence we have a reason not only why the indicative forms used in Sanskrit prohibitions are past but also why they are augmentless.⁵⁰

The mistake—as I believe it to be—made by modern scholars in assuming that the augmentless past indicative was modally used was shared, unconsciously of course, by the users of Vedic. It was natural to view an order and a prohibition as more or less parallel and equivalent; positive and negative orders must often have been associated in a single sentence. An order might use the imperative, the subjunctive, or the optative. A prohibition might use the optative with the non-modal particle *nā*, and even (in the case of one verb, *bhujema*) with the modal particle *mā*. It was not strange then that the augmentless past indicative with *mā* was misinterpreted as likewise a modal form, and came to be used outside prohibitions, as the equivalent of a subjunctive or an optative.⁵¹ But this extension was not permanent, and the original use of the form in prohibitions was the only one to survive in the later language, with the exception of a few sporadic instances.

s' applique de préférence à une action déjà commencée, à un sentiment que l'on éprouve déjà lorsque la défense est formulée.'

⁴⁸ As just noted, fn. 47.

⁴⁹ Again cf. fn. 47.

⁵⁰ In view of the interchangeability of augmented and augmentless forms to express past time, it is not surprising that analogy should introduce an occasional augmented form in prohibitions as well; this happened in the epics and later, but is met only two or three times in Vedic (Whitney 218).

⁵¹ In the same way I believe that in Latin the past indicatives of the early type referred to in fn. 42 generated the *-ā*-subjunctive. This was used as a present tense, and so eventually the *-ā*-endings came to be added to the present stem instead of to the aorist stem as was originally the case. For a more detailed discussion of the problem of the *-ā*-subjunctive, and of some of the other matters here referred to, see my monograph, *Subj. and opt.* 34–51.

LA FLEXION PRONOMINALE EN HITTITE

EMILE BENVENISTE

Collège de France

La comparaison de la flexion pronominale hittite avec celle des langues apparentées fait ressortir immédiatement deux traits distinctifs: d'une part le hittite se conforme au modèle indo-européen en attribuant des flexions différentes aux noms et aux pronoms; de l'autre, la flexion pronominale n'a en hittite à peu près rien de commun avec le système admis comme indo-européen. Cette situation mérite examen, et l'objet de ces pages est même de montrer que certaines correspondances tenues pour certaines sont fallacieuses, et que c'est en soulignant ces différences qu'on prépare une interprétation meilleure des faits hittites aussi bien que de la préhistoire des pronoms en indo-européen.

Les désinences hittites seront ici commentées selon l'importance des problèmes qu'elles offrent, et aussi bien dans les pronoms personnels que dans les démonstratifs. Ces deux classes de pronoms ont certaines désinences en commun, mais non pas les mêmes que dans les autres langues. Pour les démonstratifs, on utilisera principalement *kaš* 'hic' et *apaš* 'is', avec les formes casuelles du singulier: nom. *ka-a-aš*, *a-pa-a-aš*; acc. *ku-u-un*, *a-pi-u-un*; gén. *ki-e-el*, *a-pi-e-el*; dat. loc. *ki-e-da-ni*, *a-pi-e-da-ni*, ainsi que les formes correspondantes du pluriel.

Le NOMINATIF de ces formes n'appelle aucune remarque spéciale, puisque la désinence est identique à celle des formes nominales. Mais cette absence de distinction est elle-même intéressante. Le hittite ne connaît donc pas de nominatif pronominal à désinence zéro du type de **so*, gr. *ὁ*; le *-š* nominal a été généralisé. Un second trait négatif, également notable, et probablement lié au précédent, est que la flexion ne présente pas dans les démonstratifs une alternance de thèmes comparable à celle de **so/*to*.

Au sujet du GÉNITIF, il est bon de remarquer expressément qu'il n'y a pas trace en hittite de la désinence **-syo*, **-so* typiquement pronominal.¹ La formation caractéristique *-el* a été déjà abondamment commentée.² Ce génitif est commun au singulier et au pluriel des pronoms personnels: sg. *ammel*, *tuel*; pl. *anzel*, *šumel* (et aussi *šumenzan*³), alors que dans le reste de l'indo-européen le singulier et le pluriel ont des désinences différentes. Cela donne à penser que *-(e)l* n'était pas proprement casuel, mais plutôt adverbial. Certes l'emploi de *-l-* est bien connu dans la dérivation indo-européenne et hittite, mais sur la base de verbes ou de noms. Il est exceptionnel de le rencontrer dans les pronoms. En fait on ne le constate, comme H. Pedersen l'a indiqué,⁴ que dans la forma-

¹ Nous ne retiendrons pas la suggestion de H. Pedersen, *Hittitisch* 71, qui veut retrouver **-syo* dans *maši-* 'combien', *mašiyant* 'combien grand'.

² Bibliographie chez Sturtevant, *HG*² §199 et n. 27. Pour F. Sommer, *Hethiter und Hethitisch* 86 (1947), la finale *-el* viendrait du hatti.

³ Nous indiquons tout de suite qu'il ne sera pas question ici de l'abl. en *-az* ni du g. pl. en *-enzan*. Le problème est d'abord phonétique et concerne *z*. Nous espérons y revenir ailleurs.

⁴ Op.cit. 55.

tion de lat. *tālis*, *quālis*, gr. *τηλικος*, *πηλικος*, v.sl. *tolī* 'autant', *tolīkū kolīkū* 'aussi grand', etc. C'est là d'ailleurs une classe d'emploi assez différente et constituée seulement par des adjectifs de sens dimensionnel. Les formes de génitif hittite pourraient néanmoins s'y rattacher par le détour d'une construction prédicative d'appartenance: ceci est *ammel* 'de moi, mien'; ceci est **k^wal-* 'de telle grandeur'. De toute manière le hittite a divergé de bonne heure, car la suffixation en *-l* est certainement une des plus anciennes (cf. le nom de structure archaïque **sāwel* 'soleil' et sa flexion hétéroclitique en *l/n*). Il devient d'autant plus intéressant d'observer que le lycien a justement des adverbes en *-l* formés sur base pronominale, *ebeli* 'ici', *tehi* 'là'. H. Pedersen a bien vu que *ebeli* se compare au génitif pronominal hittite *apel* 'eius';⁵ et si le thème **to-* existait en hittite, l'adverbe lycien *tehi* y aurait pour correspondant un génitif **tel*. L'origine adverbale de la désinence hittite devient par là probable. Mais ce n'est pas seulement avec le lycien que le hittite s'accorde dans la constitution de formes pronominales en *-l*. Le lydien aussi a une désinence en *-l* pour le cas oblique des pronoms comme des noms:⁶ *esl mruλ* 'de (?à) cette stèle', et particulièrement l'oblique *bil*, probablement 'eius' (d'où l'adjectif *bilis* 'sien, propre'). Peut-être même avons-nous dans ce dernier pronom l'équivalent étymologique de lyc. *ebe*, hitt. *apa-* 'is', de sorte que *bil* répondrait exactement à lyc. *ebel-i*, hitt. *apel*. En tout cas une corrélation morphologique importante relie ces trois langues et montre un développement auquel les autres langues indo-européennes n'ont participé qu'assez faiblement. Le génitif pronominal hittite a ses attaches les plus claires dans le groupe asianique de l'indo-européen.

Aux yeux de tous les auteurs qui ont traité de la flexion hittite, l'ACCUSATIF pronominal ne pose aucune question: immédiatement et presque sans débat, *-un* (acc.sg.) a été expliqué par **-om* ou **-m*, et *-uš* (acc.pl.) par **-ons* ou **-ns*.⁷ S'il nous paraît nécessaire ici de mettre en discussion ce qui s'est imposé comme une évidence, c'est d'abord que l'accusatif nominal introduit ici une grave difficulté.

Rappelons que toutes les classes de noms de genre animé ont en hittite un accusatif sg. en *-n* précédé de la voyelle caractéristique du thème, et un accusatif pl. en *-uš*. Ainsi, en nous bornant à un exemple par classe flexionnelle: *attaš* 'père': *attan*, pl. *attuš*; *halkiš* 'blé': *halkin*, pl. *halkiuš*; *lingaiš* 'serment': *lingain*, pl. *lingauš*; *heuš* 'pluie': *heun*, pl. *heuš*; *harnauš* 'siège d'accouchement': *harnaun* (pas d'acc. pl. attesté); *kartimmiyaz* (thème en *-t-*, nom. *-at-s*) 'colère': *kartimmiyattan*, pl. *kartimmiyattuš*. Hors des noms neutres, l'opposition nom. *-š*: acc. *-m* est caractéristique de la flexion nominale, et l'accusatif pluriel en *-uš* est commun aux noms de genre animé et aux pronoms en hittite. Cette double constatation est essentielle pour la discussion des accusatifs pronominaux.

La question se pose de savoir comment expliquer historiquement la différence de vocalisme entre *-an* nominal (acc. *attan*) et *-un* pronominal (acc. *kun*, *apun*). H. Pedersen considère hitt. *-an* comme issu de **-ām* ou **-m*, et hitt. *-un* de

⁵ Pedersen, *Lykisch und Hittitisch* §30-1 (1945).

⁶ Voir les formes lydiennes chez Kahle-Sommer, *Kleinas. Forsch.* 1.46, 78.

⁷ Dans la discussion de ces formes, il a souvent été question de la désinence de 1ère sg. du prétérit *-un*. Nous laissons ce point de côté pour le traiter ailleurs.

*-om; pour -uš plur., il pense à *-ons ou *-ōns.⁸ De même pour Sommer, -uš sort de *-ons ou *-ns suivant les thèmes.⁹ Sturtevant tire -an de *-om, et -un -uš de *-n *-ns.¹⁰ Ainsi de l'avis de tous, la distinction de -an nominal et de -un pronominal refléterait une variation préhistorique entre *-om et *-n. Mais les opinions varient largement sur les correspondances phonétiques à poser: tandis que Pedersen ramène -an à *-n et -un à *-om, Sturtevant intervertit les rapports et pense que -an vient de *-om et -un de *-n. Que les deux propositions puissent être également soutenues sans que l'une réfute immédiatement l'autre est la preuve ou que le problème n'admet pas de solution ou que la solution est ailleurs. Du reste Sturtevant lui-même, le seul qui ait considéré attentivement la question, ne se sentait guère satisfait de la réponse qu'il y avait donnée. Pour lui, 'in view of the contrast between ta-an "et eum" and ku-u-un "hunc", it is safer to derive the former from -om and the latter from syllabic n'. Il reconnaissait cependant que 'the evidence of other words suggests rather that syllabic n became Hitt. an or, before s, either ant or a; perhaps we should reconstruct IH ʾn'.¹¹ Il admettait de même que r syllabique donne hitt. ur, mais que d'autres mots indiquent ar, et il essayait de les expliquer par ʾr. Son embarras se marque dans sa conclusion: 'The interpretation of Hitt. an, ar and al as containing the reduced vowel ʾ is admittedly merely a makeshift, since un and ur demonstrably represent syllabic m, n, and r. If a better solution of the problem can be found, I shall be pleased.'¹²

Il faut en effet lier les traitements de *-n et de *-r dans le même examen. Ecartant, comme Sturtevant était prêt à le faire, l'hypothèse d'une voyelle réduite, nous devons chercher comment, hors du cas en discussion, une consonne syllabique est représentée en hittite. En faveur d'un traitement ur de *-r, Sturtevant a allégué trois exemples: hurkel 'death penalty': harkzi 'is destroyed': irl. orgaim 'ich schlage, erschlage'; andurza 'within' < *an-dur: osco-umbr. anter, lat. inter; hurtaiš 'curse': gr. εἶπω 'say' (pf. εἶρηκα), got. waird 'word'.

Mais diverses objections éliminent ces exemples un à un: hurkel signifie non 'peine de mort', mais 'abomination, atrocité', et n'a sans doute rien de commun avec hark- 'disparaître, s'anéantir'; andurza n'a pas dû sembler très convaincant à Sturtevant lui-même pour qu'il en ait donné par mégarde, à quelques pages de distance, deux explications contradictoires: §64 'andurza "within" < *an-dur: Lat. inter', mais §84 'andurza "within" < *ʾn-dʾur-ts "indoors": θῆρα, Lat. fores';¹³ hurtaiš 'malédiction', s'il s'apparente à gr. εἶπω, got. wairds, doit reposer sur *hwerdh-, dont hitt. hurd- sera la degré réduit; donc de -u- est ici en alternance avec *-we- et ne concerne en rien le traitement de r, qui est simplement consonantique.

Il n'existe donc aucune preuve d'un -ur hittite issu de *-r. Tous les faits clairs enseignent au contraire que *-r est représenté en hittite par ar: cf. arnuzi: skr.

⁸ Pedersen, *Hittitisch* 21, 49, 90.

⁹ Sommer, *Hethiter u. Hethitisch* 48.

¹⁰ HG² §63 seq.

¹¹ HG² §63.

¹² HG² §65.

¹³ Cf. maintenant le même embarras au sujet de u chez Friedrich, *Hethitisches Wörterbuch* 24 (1952): 'andurza zu lat. inter, ai. antar mit u wie in lit. kur "wo?"'

ṛnoti; *arškizi* : skr. *ṛcchati*; *parku* < **bhrghu* : arm. *barjr* 'haut' (thème en -u < **bhrghu*-), *ešhar* : skr. *aṣṭ-k*, gr. *ἔαρ*; *kardi* dat. : skr. *hṛd*-, gr. *καρδία*. Comme par ailleurs hitt. *ar* peut représenter aussi **or* (*daru* : *δῶρυ*), on conclura que i.e. **or* et **r* aboutissent également à hitt. *ar*.

Il en est de même pour la nasale sonante. Nous avons des faits probants comme : *panku* < **bhṇghu* : gr. *παχὺς*, skr. *bahu*-; *dankui* 'sombre, noir' < **dhṇghw*- : all. *dunkel*; *anzaš* 'nous' : all. *uns*; -*anki* suffixe multiplicatif : gr. -*ακι*. Il est par ailleurs établi que hitt. -*an* continue un *-*om* ancien dans l'acc.sg. des noms en -*a*- comme *atta*- 'père', acc. *attan* ou dans le nom. acc. d'un neutre comme *yugan* 'joug' < **yugom*. Les formes du démonstratif enclitique nom. -*aš*, acc. -*an* remontent à **os* : **om*. On doit donc admettre que *-*om* et *-*ṇi* donnent également -*an* en hittite. Il y a symétrie dans les deux traitements : **or* et **r* > -*ar*; *-*om* et *-*ṇi* > -*an*.

Dès lors on ne voit plus aucune possibilité de ramener la désinence d'accusatif pronominal -*un* soit à *-*om*, soit à *-*ṇi*. Les deux solutions sont pareillement exclues. Et ce sont les deux seules qu'on ait envisagé jusqu'ici.

Pour retrouver le prototype de -*un*, il faudra nécessairement partir de cette constatation qu'une sonante est impossible, et admettre que la finale à restituer sera de la forme 'voyelle + *m*'. Cela revient à chercher quelle voyelle indo-européenne peut être représentée par hittite *u*. La réponse est aisée. Un *u* hittite répond toujours et seulement à i.e. **u*. Outre *daru*, *parku*-, *panku*-, *yugan* déjà cités, et qui suffiraient à justifier la correspondance, mentionnons *nu* : gr. *νυ*; (*ki*)*nun* : gr. *νύν*; *genu* : lat. *genu*; *luk*- : lat. *luc*-; *kuwapi* 'quelque part' : lat. (*ne*)*cubi*.

La conclusion est inévitable : l'accusatif sg. pronominal -*un* suppose non *-*om* ni *-*ṇi*, mais seulement *-*um*. Une pareille désinence peut nous déconcerter. Que nous ne puissions lui trouver de place dans les cadres traditionnels de la morphologie indo-européenne, alors qu'elle nous est imposée par des correspondances certaines, prouve seulement que nous ne connaissons pas encore toute cette morphologie, et que les faits ont pu être encore plus complexes que nous ne le pensions. Mais heureusement nous ne sommes pas réduits à accepter seulement ce type insolite d'accusatif, nous pouvons le légitimer hors du hittite. Cette désinence *-*un* permet d'élucider une forme pronominale restée énigmatique en sanskrit : l'accusatif sg. *amúm*.

Nous touchons ici à l'une des situations les plus confuses de la flexion pronominale indienne. La forme *amúm* tient la place d'accusatif sg. dans le paradigme hétéroclite du démonstratif *aśu* 'hic' (av. *hau*, v.p. *haw*) qui a pour neutre *adāḥ*.¹⁴ Sur le thème *amu*- se constituent, outre *amúm*, les cas obliques du singulier et les formes du pluriel : gén. *amūšya*, dat. *amūšmai*, abl. *amūšmāt*, loc. *amūšmīn*, etc. Comme Wackernagel-Debrunner l'ont observé,¹⁵ ce thème *amu*-généralisé dans une grande partie de la flexion procède de l'acc. *amúm*, qui a été le point de départ de cette normalisation partielle. Mais comment expliquer un accusatif *amúm*? Ces auteurs sont, comme leurs devanciers, contraints à

¹⁴ L'exposé le plus détaillé du problème est chez Wackernagel-Debrunner, *Altind. Gr.* 3.528 seq.

¹⁵ Op.cit. 530.

une analyse bien artificielle: *amum* serait formé de *am*, acc.sg. du pronom *a-*, augmenté de la particule déictique *u*, augmenté encore d'un autre *-m* accusatif. Ces complications peuvent être abandonnées pour une analyse plus directe et simple. L'accusatif *amum* doit se coupler avec un autre démonstratif, isolé lui aussi, représenté par le nominatif sg. *amas*, et attesté dans la vieille formule matrimoniale *ámo 'hám asmi, sá twám* 'je suis lui, tu es elle' (AV). Une variante *amúham* (TB) souligne la relation encore perçue entre *amas* et *amum*. D'ailleurs le thème *ama-* est maintenant confirmé par le témoignage indépendant de l'iranien, qui fournit l'adverbe v.p. *amata* 'de là'.

Il reste maintenant à poser ces deux formes comme les membres survivants d'un vieux paradigme disloqué: nom. *amas* et acc. *amum*. Alors nous obtenons un schème d'oppositions exactement symétrique en indien et en hittite:

Skr. nom. <i>amas</i>	< * <i>emos</i>	: acc. <i>amum</i>	< * <i>emum</i>
Hitt. nom. <i>apaš</i>	< * <i>ebhos</i>	: acc. <i>apun</i>	< * <i>ebhum</i>
nom. <i>kaš</i>	< * <i>kos</i>	: acc. <i>kun</i>	< * <i>kum</i> .

Cet accord paraît garantir l'authenticité et l'antiquité d'un type flexionnel encore inconnu. Il semble que la désinence **-om*, seule attestée dans les autres langues à l'acc.sg. des pronoms et des noms, ait succédé, dans la flexion des démonstratifs, à un plus ancien **-um*, que le hittite préserve dans sa fonction régulière et dont un vestige isolé, la forme *amum*, a été en indien incorporé à un paradigme composite. L'interprétation de l'acc.sg. enclitique hitt. *-an* reste incertaine: on peut penser à une extension de *-an* nominal, ou à une coexistence de *-an* et *-un* selon les classes pronominales. L'acc.pl. est *-aš* et *-uš* à la fois.

Pour rendre compte de l'accusatif pluriel pronominal et nominal en *-uš*, nous n'avons plus besoin de **-ons* ni de **-ns*. La caractéristique *-u-* de l'accusatif sg. a été transférée au pluriel et munie du *-š* qui signale en hittite tous les cas du pluriel hors des neutres, à l'exception du gén. en *-enzan*. De là *kuš*, *apuš* aussi bien que *attuš*. L'accusatif pluriel pronominal a été étendu aux noms.

Ces considérations donnent le moyen de simplifier l'explication historique des pronoms personnels. La distribution des formes de nominatif et d'accusatif en hittite s'écarte sensiblement du modèle indo-européen. On a les oppositions suivantes: (1) nom. *u-ug* /ug/ 'ego', acc. *am-mu-uk* /amug/ 'me'; (2) nom. *zi-ig* /teg/ 'tu', acc. *tu-uk* /tug/ 'te'. Il est admis depuis Sturtevant¹⁶ que la voyelle de *tug*, qui est originale (cf. gr. *σὺ*, got. *þu-k*), est passée à *amug*, lequel à son tour a influencé *ug*. On peut maintenant répartir autrement ces deux extensions analogiques. A la 1^{ère} sg. l'opposition ancienne en hittite était vraisemblablement **eg* 'ego' : *amug* 'me'. Cet **eg* (= got. *ik*) a pris le vocalisme de *amug*, auquel la différence de thème suffisait à l'opposer; de là *ug* : *amug*. Cet accusatif *amug* est évidemment apparenté à la forme enclitique *-mu* 'me, mihi'. Ce n'est sans doute pas un hasard si nous retrouvons à l'accusatif du pronom personnel, qui n'a jamais comporté de *-m*, le même vocalisme *u* qui caractérise l'accusatif des démonstratifs **-u-m* avec adjonction régulière de *-m*. Il semble que la distinction de timbre *e* : *u* ait eu valeur morphologique, dans un état très ancien de la flexion pronominale indo-européenne, pour réaliser l'opposition

¹⁶ HG² §170b.

du cas sujet au cas objet. Dès lors ce principe peut aussi expliquer la relation nom. *teg* : acc. *tug*, exactement symétrique de **eg* : *amug*, par une extension de la 1ère à la 2e personne et non en sens inverse. Ici à la différence de l'autre pronom, la distinction des timbres a été maintenue puisqu'elle jouait seule un rôle grammatical. Il est plus difficile d'apprécier la relation entre les formes hittites ainsi définies et celles des autres langues où, en sens contraire, *u* est la marque du nominatif et *e* de l'accusatif à la 2e personne (gr. *τυ/σε*). Quand on connaîtra plus complètement la morphologie pronominale des langues d'Asie-Mineure apparentées au hittite,¹⁷ quelque lumière se fera sans doute sur la préhistoire des pronoms personnels dans ce groupe dialectal.

Le DATIF singulier des démonstratifs présente les deux désinences *-edi* et *-edani*, sans compter une forme *-eda* plus rare. Ces formes servent pour le datif et le locatif, et au pluriel cette double fonction est remplie par la forme unique *-edaš*. Certains pronoms étendent ce procédé de formation à l'ablatif: ainsi les pronoms personnels *ammedaz*, *tuedaz*; mais les démonstratifs le limitent au datif-locatif et font à l'ablatif *kez*, *apez*. Pedersen indique avec raison que cette désinence *-edi*, *-edani* doit procéder d'une finale adverbale ancienne, mais il n'en fournit pas d'interprétation. Pour Sturtevant, le dat. *apeda* suggère une comparaison avec les adverbes skr. *yadā*, *tadā*.¹⁸ Ce rapprochement nous paraît d'autant plus probable qu'il y a en indien, dans les démonstratifs, échange fréquent entre les valeurs casuelles et adverbiales; cf. *idam* 'ceci' et 'ici'; *adaḥ* 'cela' et 'là'. Du fait que la formation de skr. *yadā*, *tadā* est indo-iranienne (gāth. *kadā*, *yadā*, av. *kaša*, *taša*, *iša*), l'antiquité de la désinence hittite *apeda* 'huic' est garantie par un groupe de formes certainement héritées,¹⁹ et dont le nombre, comme on va voir, peut s'accroître encore.

Aucune explication n'a encore été donnée du datif en *-edani*. La question de savoir si c'est un élargissement secondaire de *-eda* ou au contraire une désinence ancienne ne peut être tranchée que par un rapprochement probant. Or nous croyons pouvoir l'identifier avec une finale adverbale du védique: celle des adverbes temporels et locaux *idānīm* 'maintenant', *tadānīm* 'alors', *viśvadānīm* 'en tout temps', bien attestés dans les hymnes et en prose: *kvedānīm sūryaḥ* 'où est maintenant le soleil?' (I 35.7); *nāsad āsīn nō sād āsīt tadānīm* 'il n'y avait ni être ni non-être en ce temps' (X 129.1); *addhī tñam aghnye viśvadānīm* 'mange l'herbe en tout temps, ô vache!' (I 164.40). Ces formes, d'un type insolite et inexplicable en indien, sont munies de la finale *īm* propre à quelques anciens adverbes indo-iraniens: ved. *tūṣṇīm* 'silencieusement' (cf. av. *tuṣni-šad* 'assis en silence', *tuṣnā-matay-* n.pr. fém.); av. *xrūmīm* 'cruellement', cf. *xrūma-* 'cruel'. Il devient possible ainsi de ramener *idānīm* *tadānīm* à des adverbes **idāni* **tadāni* comparables à hitt. *kedani* *apedani* et qui, comme ces formes hittites, ont une fonction de locatif. Quant aux formes 'courtes' de dat. loc. pronominal, hitt. *edi*, *kedi*, *apedi*, c'est aussi dans les adverbes indo-iraniens qu'elles trouvent

¹⁷ Si intéressantes que soient les formes h.hiér. *amu* 'ego', lyc. *emu* 'ego, me', lyd. *emu* 'ego', on n'en pourra apprécier la portée que quand le système pronominal de ces langues sera établi dans son ensemble.

¹⁸ *HG*² §198.

¹⁹ En h.hiér. le dat. loc. sg. semble être 'apata: Oney, *JCS* 5.62.

des parallèles: *kedi* de *ka-* se repiera à i.ir. *yadi* 'si' de *ya-*; et si véd. *tādītnā* 'alors, en ce temps' contient, comme il est probable, un ancien **tādi* suffixé par *-tna-* temporel, il fournit un second exemple du type adverbial sur base pronominale. C'est toute une série de correspondances qui se dévoilent ainsi du hittite à l'indo-iranien dans la morphologie pronominale, situant et coordonnant en une perspective plus ancienne nombre de traits isolés et anomaux.

Un dernier problème à considérer sera la curieuse particularité d'une forme commune aux 2e et 3e personnes du pluriel dans les formes pronominales et possessives enclitiques du hittite. En forme libre, le pronom de 2e pl. est *šumeš* 'vous'. Mais la forme enclitique *-šmaš* de datif et d'accusatif signifie aussi bien '(à) vous' que '(à) eux'. De même l'enclitique possessif est commun aux deux personnes: sg. n. *-šmit* 'votre' et 'leur'; pl. nom. animé *-šmeš* 'vos', dat. n. *-šmaš* 'à leurs'. Ce trait insolite est simplement enregistré par Sturtevant: '*šmaš* and the corresponding possessive pronoun are alone among Hittite and early IE pronouns in referring alike to the second and third persons'.²⁰ Apparemment il n'a pas été convaincu par la tentative d'explication avancée par Pedersen,²¹ selon qui *-ši* sortirait d'une forme d'enclise répondant au dat. f. sg. skr. *asyai*, et *-šmaš* serait la pluralisation d'une forme répondant au dat. m. sg. skr. *asmai*.

Hitt. *šumeš* 'vous' ressemble à la forme persane du même pronom *šumā*. Et comme *šumā*, à travers phl. *šmāk*, remonte à *xšmāka-* bien établi en avestique, on est tenté d'expliquer hitt. *šumeš* par une forme comparable à av. *xšma-*.²² Il faut se garder de cette méprise. L'avestique a les deux séries *xšmaṭ xšmaibyā xšmākəm* et *yūšmaṭ yūšmaibyā yūšmākəm*, mais l'initiale de *xšma-* suppose **šma-* qui ne peut être phonétique et provient de *yūšma-* d'où il a été extrait. On ne peut donc rien conclure du pronom avestique, qui est un doublet secondaire, pour la restitution de la forme hittite. Il faut partir, pour le hittite comme pour l'ensemble des autres langues, de i.e. **usme* (plus probable que **yusme*)²³ où **us-* alterne avec **wes* (skr. *vaḥ* etc.). C'est comme continuation de **usme* que hitt. *šumeš* (pluralisé en *-š* d'après *weš* 'nous', got. *weis*) s'explique au mieux, en admettant une métathèse dans la première syllabe. Et, au point de vue indo-européen, **usme* 'vous' et **nsme* 'nous' se définissent par l'addition de *-m-*, non *-sm-*, à **wes* et **nes*.

Prenant maintenant les formes de 3e pl. hittite, on observera, en les confrontant à celles du singulier du possessif enclitique, qu'elles se caractérisent par l'insertion d'un *-m-* avant la désinence:

sg. acc. *-š-an*; dat. *-š-i*; instr. *-š-it*
pl. acc. *-š-m-an*; dat. *-š-m-i*; instr. *-š-m-it*.

Le *-m-* a ici un rôle pluralisant dont nous ne connaissons pas d'exemple ailleurs. Peut-être vient-il du pronom personnel pluriel. On ne saurait donc rien affirmer au sujet d'une relation avec le *-m-* qui pourrait être dégagé des cas obliques du

²⁰ HG² §175.

²¹ Hittitisch 58.

²² Ainsi Sturtevant §175.

²³ Meillet, BSL 23.76.

démonstratif au singulier, tels que skr. *tasmai*, ombr. *esmei*, got. *imma*, etc. si l'on analyse *-sm-* en *-s-* (cf. *-s-ya*, f. *-s-yās*) + *-m-* (cf. sl. *to-m-u* dat. etc.).

A les considérer ainsi dans leurs formes respectives, ces pronoms de 3e pl. apparaissent comme différents de *šumeš* 'vous'. Le possessif enclitique *-š(u)meš* 'vos' (plur. de 'votre') va naturellement avec *šumeš* 'vous'. Mais *-š(u)meš* 'leur' représente *-š-m-eš*, pluralisation de **-šmaš* (ou **-šmiš*) 'leur', comme *-šeš* 'ses' pluralise *-šiš* 'son'. Entre *-šmeš* 'vos' et *-šmeš* 'leurs' il n'y a plus identité, mais simple coïncidence à partir d'origines distinctes. Rien n'assure même que les deux *-šmeš* aient été phonétiquement pareils. Il vaudrait la peine d'étudier de plus près les graphies multiples, *-šum-mi-it*, *-ša-mi-it*, *-ši-mi-it*, *-(i)š-mi-it*, etc., qu'on s'est peut-être trop hâté de déclarer équivalentes. Un relevé détaillé de tous les exemples en contexte sûr serait souhaitable pour qu'on pût décider si les deux séries que nous distinguons étaient ou non restées distinctes. Une confusion totale serait a priori peu vraisemblable.

Ces remarques suffiront à montrer que la flexion pronominale du hittite ouvre des perspectives imprévues sur la préhistoire morphologique de l'indo-européen et modifie la position traditionnelle de maints problèmes. On peut espérer que le déchiffrement des autres langues indo-européennes d'Asie-Mineure précisera quelques unes des solutions suggérées ici et contribuera à élucider cette chronologie de l'indo-européen à laquelle le hittite a apporté tant de données précieuses.

THE THEOPHOROUS ELEMENTS OF THE ANATOLIAN PROPER NAMES FROM CAPPADOCIA

ALBRECHT GOETZE

Yale University

Among the Anatolian proper names transmitted to us in the Old Assyrian tablets from Cappadocia, those in *-uman* and *-l(i)ka*, in *-aḥšu*, *-aḥšu-šar*, and *-aššu*, and finally those in *-iyat/iet* are particularly characteristic. Their analysis, undertaken elsewhere,¹ has shown that they contain the following identifiable theophorous elements: *Aššiyat*, *Ḫalki*, *Ilaliya*, *Inar(a)*, *Išput*, *Perwa*, *Tarawa*, and possibly *Šiwat*. The question arises whether these gods can be ascribed to a definite ethnic and linguistic layer of early Anatolia, such as is known to us from the Boğazköy texts.

Since the very beginning of Hittite studies it has been recognized that the Hittites worshiped gods of different origin, and that they felt it necessary to address foreign gods in their native languages.² These languages are frequently defined. To some deities songs are to be recited *ḫattili* 'in Ḫattic', to some of them *ḫurlili* (*hurili*) 'in the fashion of a Hurrian', and to still others *palaumnili* 'in the fashion of a man of Palā' or *luwili* 'in the fashion of a Luwian'. In certain cases, finally, the 'singer of Kaniš' (^{LÜ}NAR ^{URU}*Ka-ni-iš/eš*)³ sings. For identifying the language to which modern writers apply the term 'Hittite', the texts themselves say *našili*, *nešili*, or *nešumnili* 'in the fashion of (the city) Neša' or 'in the fashion of a man of Neša'.⁴

¹ In an article to appear later in this journal.

The following abbreviations in the present article may be unfamiliar to readers of *Language*:

(1) Periodicals: *Bibl. Or.* = *Bibliotheca orientalis*, *JCS* = *Journal of cuneiform studies*, *JHS* = *Journal of Hellenic studies*, *JKF* = *Jahrbuch für kleinasiatische Forschung*, *OLZ* = *Orientalistische Literatur Zeitung*, *RA* = *Revue d'assyriologie*, *RHA* = *Revue hittite et asianique*, *ZA* = *Zeitschrift für Assyriologie*, *ZDMG* = *Zeitschrift der Deutschen Morgenländischen Gesellschaft*.

(2) Publications of Boğazköy texts: *ABOT* = Kemal Balkan, Ankara Arkeoloji Müzesinde bulunan Boğazköy tabletleri; *HT* = Hittite Texts in the Cuneiform Character from Tablets in the British Museum; *IBOT* = Bozkurt-Çiğ-Güterbock, Istanbul Arkeoloji Müzelerinde bulunan Boğazköy tabletleri; *KBo* = Keilschrifttexte aus Boghazköi; *KUB* = Keilschrifturkunden aus Boghazköi (usually quoted by number of volume only). Furthermore *Bo* (with following number) = unpublished Boğazköy tablet (quoted by museum number).

(3) Other publications of cuneiform tablets: *TCL* = Musée du Louvre, Textes cunéiformes.

² E. Forrer, *Die acht Sprachen der Boghazköi-Inschriften* (SPAW phil.-hist. Klasse 1919.1029 ff.); F. Hrozný, *Über die Völker und Sprachen des alten Chatti-Landes* (*Boghazköy-Studien* 5, 1920).

³ Only recently the adverb *kanišumnili* 'in the fashion of a man of Kaniš' has emerged: Otten, *JKF* 2.67 (1951) (Bo 6222).

⁴ The reasoning of H. T. Bossert, *Ein Hethitisches Königssiegel* 16 ff. (1944), leading to the conclusion that 'Neshite' is not Hittite, is invalid. Both in *KBo* V 11 I 3 f. and in *Bo* 2002—now published as *IBOT* I 36—III 64 the words introduced by the adverb *nešili* are

This feature enables us, as a rule, to distribute the gods venerated by the Hittites among the various layers of which the population of Hittite Anatolia consisted. It was first utilized by E. Forrer (ZDMG 76. 192 ff. [1922]), and more recently by H. T. Bossert in his book *Ein hethitisches Königssiegel* (1944). It is the thesis of this article that the deities listed above are attributable to the Kanishites, who are closely related to the Luwians. To prove our thesis we have to look into the god lists to be culled from ritual texts. The Kanishite lists are as follows:

(1) Sacrifices for⁵ *Pi-ir-wa* ^PSAL.LUG[AL] ^PAš-ka-še-pa ^PVII.VII-BI ^PŠu-wa-li-ia-at[-ti] DINGIR.SAL.MEŠ-ia ^PŠi-wa-at-ti ^PHa-ša-am-me-l[i] DINGIR.MEŠ ^{URU}Ka-ni-iš Bo 2597 + Bo 2659 II 4 ff. quoted by Otten, JKF 2.73 fn. 29.

(2) Bread offerings to⁶ ^PA-āš-ga-še-pa, ^PSAL.LUGAL, ^PPi-ir-wa, ^PVII.VII-BI, DINGIR.LŪ.MEŠ(-aš) ^PMa-li-ia(-aš) DINGIR.LŪ.MEŠ ^PKAL-aš, DINGIR.LŪ.MEŠ(-aš) ^PUD.KAM-aš, DINGIR.MEŠ ^{URU}Ka-ni-iš, ^PHa-ša-am-me-li, ^PHi-la-aš-ši, ^PU.GUR, ^{ÍD}.MEŠ⁷ KUB II 13 III 16 - IV 3 (cf. IV 12-28).

(3)⁸ ^PPi-ir-wa-an, ^PAš-ka-ši-pa-an, ^PSAL.LUGAL-ri, ^PMa-li-ia-an / DINGIR.-MEŠ A-BI ^{hu-u-ma-an-te-eš} DINGIR.MEŠ ^{URU}Ha-at-ti ^{hu-u-ma-an-te-eš} DINGIR.MEŠ KUR-TI ^{hu-u-ma-an-te-eš}, both groups sung by the ^{LŪ}NAR ^{URU}Kaneš KBo IV 13 VI 9 ff.

(4) Animals are sacrificed to⁹ gods who are enumerated in a rising order beginning with the lower ranks: *a-aš-ša-u-i* MUL-i, ^PMAH.M[EŠ] [^PGul-ša-aš¹⁰], A-NA ^PKAL ^{KUB}kur-ša-aš EN-i š[al-la-i¹¹], ... , ^PUD-aš ^PVII.VII-BI, ^PX-aš ^P[VII.VII-BI], ^PKAL-aš ^PVII.VII-BI, ^PAš-ka-ši-[pa], ^PSAL.LUGAL-ri, ^PPi-ir-wa, ŠA DINGIR.MEŠ LŪ.MEŠ ^PMa-li-ia KBo IV 13 I 10 ff.

(5) We may add here already the sequence of gods—apparently proceeding from the lower to the higher rank—which is contained in the mythological text KBo III 8 III 4 ff.: *IŠTAR, Maliyaš, Pirwaš, Kamrušepaš*.

Before subjecting the gods enumerated in these lists to a more detailed treatment we must mention the fact that the sequence *Pirwaš—Aškašipaš*—[^PSAL.-LUGAL] also figures in XXV 32 I 11 and *Pirwaš*—^PSAL.LUGAL also in IBoT II 131 rev. 4.

In the god lists that form an integral part of Hittite treaties occurs a pertinent

in Hittite (*halugaš* 'message' and *tapuša* 'to the side'). Contrast with ^{URU}ni-ši-li *hal'-za'-i'* *ta-pu-ša* in the same text I 65 *nu ha-at-ti-i-li ta-ha-ia hal-za-i*, and compare Goetze, *Tunnawi* 34. In the 'second Arzawa letter' (VBoT 2) the scribe in Egypt is asked to reply in Neshite, i.e. 'Hittite', since that language caused him less difficulty than Luwian and since the writer would presumably have had trouble in understanding Akkadian.

⁵ The names are to be understood as datives.

⁶ Again datives.

⁷ The 'singer of Kaniš' is not mentioned here, but the occurrence of the 'gods of Kaniš' and the comparison with list 1 makes the attribution certain.

⁸ Part of a larger system.

⁹ Part of a larger system; all names are in the dative.

¹⁰ This is suggested by the preceding ^PMAH.MEŠ; see A. Goetze, *Tunnawi* 55 ff.

¹¹ Restoration doubtful.

group, not attributed to a definite ethnic element. This group varies in the following way:

^d <i>Ištar multarrihu</i>	^d <i>Aškašipa</i>	^d <i>Nisaba</i> ¹²	(Mattiwaza).
^p <i>Pirwaš</i>			(Man. ¹³ and perhaps Al.).
^p <i>Pirwaš</i>	SAL.LUGAL-aš	SAL.LUGAL[...]	ABOT 56 II 14.
^p [<i>Pirwaš?</i>]	[^p A]škašepan	SAL.LUGAL-an	XXVI 11 I 124.
^p [<i>Pi-ir-w</i>]a[-aš]	[^p A]škašepaš	SAL.LUGAL-aš	XXVI 39 IV 14.
	^p NIN.GAL	^p [SAL.LUGA]L	KBo IV 10 obv. 56.

Let us now take up the gods of the Kanishite lists in order.

Pirwaš (all 5 lists) has been the subject of an exhaustive study by Otten in *JKF* 2.62-73 (1951), who finds that *Pirwaš* is a warlike god to whom the horse is sacred. Although in his cult the 'singer of Kaniš' sings (other passages are I 17 VI 41 and X 58 V 1 ff.), Otten, like Laroche (*Recherches* 87), maintains that *Pirwaš* cannot be assigned to a definite layer of the Anatolian population. This seems to me hypercritical; there is no valid reason to deny his Kanishite character. In fact he is consistently placed at the head of the Kanishite pantheon.

The spellings ^pSAL.LUGAL (lists 1 and 2) and ^pSAL.LUGAL-ri (lists 3 and 4) present a problem of identification. The ideogram means 'Queen'. The complement -ri, which in list 3 is attached to the accusative, gives us our first clue to the reading of the name. Furthermore it cannot escape notice that in list 5 *Kamrušipaš* appears instead. Now, in the Hittic-Hittite bilingual text KUB XXVIII 4, Hitt. *Kamrušipaš* corresponds to Hittic *Kataḫziburi*.¹⁴ This name seems to be meant in our lists. The same is probably true of the treaties where SAL.LUGAL carries the complements -aš (nom.) and -an (acc.), unless her name is replaced there by an epithet. In the treaties the determinative for 'deity' is lacking, a feature that is repeated with (the entirely different) SAL.LUGAL ^a*Ka-ta-pa*, e.g. KBo I 1 rev. 49.

Kamrušipaš is the patroness of healing and of magic¹⁵ and is comparable to the Gula (lit. 'the great one') of the Mesopotamians. She resides in the 'Deep', and the 'great Sea' (*šalliš arunaš*) is her child (XVII 8 IV 20). The Hittic name *Kataḫ-ziburi* consists of the well-known noun 'queen' with a qualifying element, most probably an adjective. It may well be that *ziburi* is Hittic for 'great'.¹⁶

The next deity is *Aškašipaš* (lists 1, 2, 3, 4). List 5, where *IŠTAR* seems to occur instead, suggests that *Aškašipaš* is female and an *Ištar*-like figure.

¹² Nisaba is not only the Sumerian goddess of grain, but also the scholar among the gods. It is therefore no surprise that she seems to substitute here for the goddess of healing *Kamrušipaš*.

¹³ In the two versions (XIX 50 IV 10 and XXVI 36 IV 5) the continuation varies. It may be quoted here since it includes—at least in the former—SAL.LUGAL.GAL. However, the epithet is here attributed to *Ḫebat*. With ^p*Ḫé-bat* SAL.LUGAL.GAL [^u*U-da*] and ^p*Ḫé-bat* [^e-]b[at] [^u*U-da*] (before ^u*U-da* the words ^p*IŠTAR ŠA* seem erased) compare ^a*Ḫé-bat* [*šarr*]at *šamē* Tette IV 25 f. and ^a*Ḫé-bat* *bēlet ša-me-e* ^a*Ḫé-bat* ^a*Ḫal-pa* ^a*Ḫé-bat* ^a*U-da* ^a*Ḫé-bat* ^a*Ki-iz-zu-at-ni* Matt. rev. 46 f.

¹⁴ Restore KUB XXVII 5 obv. 17b from the duplicate 4 obv. 15b.

¹⁵ Götze, *Kleinasien* 136.

¹⁶ Cf. the remarks of Laroche in *Rech.* 39, and also ^u*zi-bu-ri-aš* XXX 69 1. Is it a mere accident that GAL 'cup' appears more fully as GAL-ri (see von Brandenstein, *Bildbeschreibungen* 28)?

The formation in *-šipaš*,¹⁷ common to both *Aška-šipaš* and *Kamru-šipaš*, has its significance for our problem. The remark is in order that the suffix must stem from one of the Anatolian languages. This is proved by the pair *te/akan* 'earth' and *daganzipaš* 'earth, ground', of which *tekan* corresponds with Toch. *tkam*, i.e. is of Indo-European origin. In fact most of the words and names containing the suffix are derived from known words, and it can hardly be doubtful that, just as *^DIšpanz(aš)ipaš*¹⁸ is the demon of the night (*išpant-*), *^DAška-šipaš* is that of the gate (*aška-*)¹⁹ and *^DKamru-šipaš* (*Kammaru-šipaš*²⁰) that of the *kam(ma)-ru*.²¹ Besides *^DHuwariyanzipaš* (see below) we have *^DHuwariyaš* (Bo 492 7). It seems also clear now that *^DHilanzipaš*²² is identical with *^DHilaššiš*,²³ and *^DHantašipaš*²⁴ with *^DHantiyaššiš*,²⁵ it is no accident that the god *Šuwanzipaš* is worshipped in the town *Šuwanzana*.²⁶ In other words, the suffix creates, in one of the Anatolian languages, adjectives of appurtenance. This is certainly neither Hittite nor Luwian; there *-šiš* is used for the purpose. *Hilanzipaš* suggests Palaic, but it might also be Kanishite. The latter assumption seems to be best in 'Kanishite' lists.

The deity *^DVII.VII-BI* (lists 1 and 2), in Sumerian *imina-bi* and in Akkadian *Sebettum* 'the seven, the Heptad', has usually been explained as 'the Pleiades'. For Hittite the reading *^DŠepittaš* has been proposed on the basis of IBoT I 10 III 13, 17, where a deity of this name follows after *^DVII.VII-BI* (ll. 5, 9) in an analogous context. The identification seems doubtful.²⁷

The interpretation as 'the Pleiades' certainly does not hold good. Belonging to the circle of the healing goddess, 'the Seven' should be identified with the seven children of *En-me-šarra*, who are healing gods like *Kamrušipa*.²⁸

In list 4 'the Seven' occur only in the combinations

^DUD-aš ^DVII.VII-BI 'the Heptad of the (Lucky) Day',²⁹

^DX-aš ^D[VII.VII-BI] 'the Heptad of the Stormgod',³⁰

^DKAL-aš ^DVII.VII-BI 'the Heptad of Inar'.³¹

¹⁷ See Laroche, *RHA* 7. 3 ff. (1946); *Rech.* 67 f. New examples are *^{HUR-SAG}Hallalazipa* (von Brandenstein, *Bildbeschreibungen* 45) and *^DHantašipa* (see fn. 22).

¹⁸ *^DIšpant-šepaš*; XX 24 III 2 heading a long series which ends in the 'Körperteil-Gott-heiten' in *-šaš*.

¹⁹ See J. Friedrich, *Hethitisches Wörterbuch* 36 (1952), with older literature.

²⁰ XVII 10 II 35; XXXIV 63 15.

²¹ If identical with *kammara-* its meaning is possibly 'mist'.

²² II 4 III 28; Bo 84 (ZA NF 14 121) 23; Bo 1551/c etc. (*ibid.* 139 ff.) rev. 17; HT 26, 5; IBoT II 86 always in a Palaic milieu.

²³ See below 270.

²⁴ Bo 2416 etc.; see *RA* 45.185 (1951) and *Bibl. Or.* 8.226 (1951).

²⁵ See below 274.

²⁶ KUB VI 45 II 22 f.

²⁷ *JCS* 1.89 (1947) against Güterbock in the preface of IBoT I. It seems that KBo II 16 offers *^DVII.VII-ru*. Cf. *^DŠi-bu-ru-ú* X 89 II 15; *^DŠe-bu-ru-ú* XX 19 III 5??

²⁸ Thureau-Dangin, *RA* 16.144 ff. (1919); *Rituels accadiens* 14 ff., 24 ff. (1921); Jean, *RA* 21.97 ff. (1924); H. and J. Lewy, *Hebrew Union College annual* 17.37 ff. (1943).

²⁹ It is difficult to decide whether the sign 'ud' should be read UTU 'sun' or UD 'day'. Decision in favor of the latter is made here in view of UD.KAM-aš in the apparent parallel to be quoted presently.

³⁰ Cf. *ŠA ^DX ^DVII.VII-BI* VII 13 25; XXVIII 4 rev. 17.

³¹ The equation *^DKAL = ^DInar(aš)*, first assumed by E. Forrer (*Forschungen* 1.10 fn. 1)

They recall from list 2 the other combinations (following ^DVII.VII-BI)

DINGIR.LŪ.MEŠ ^DMa-li-ia(-aš) 'the gods of Maliyaš',

DINGIR.LŪ.MEŠ ^DKAL-aš 'the gods of Inar',

DINGIR.LŪ.MEŠ ^DUD.KAM-aš 'the gods of the (Lucky) Day'.

Note also that list 4 closes with

ŠA DINGIR.MEŠ LŪ.MEŠ ^DMa-li-ia 'Maliyaš of the gods'.

One gains the impression that the 'gods' referred to in the last two places are 'the Seven' of the first. In any case we will have to assume that there existed different groups of gods called 'Heptad, the Seven', associated with various gods.

With respect to *Maliyaš* (who, by him- or herself, occurs only in list 5), it should furthermore be remarked that the list IBoT II 19 (see below) contains the entry [DINGIR.LŪ.]MEŠ ŠA ^DÉ.A 'the gods of Ea'. One may at least ask whether ^DÉ.A does not serve in that passage as an ideogram for ^DMaliyaš. The latter is also marked as a river,³² Ea on the other hand is the lord of the "Deep" with whom magical wisdom originates. However, Maliyaš is not a major deity as Ea would have to be. Maliyaš also occurs among the gods to whom the king and the queen 'drink' in the large ritual represented by the three duplicates XXV 2 + 6, XX 28, and XI 16, provided XX 67 (+IBoT II 77 and parallel to X 69 V) also belongs to that text. The deity is mentioned XX 67 VI 9, but the exact place in the system of the gods which it enumerates cannot be determined as yet. It can merely be observed that Maliyaš appears rather late in the list. Compare furthermore ABoT 14 IV 6 (after *Hašammeliš* I 10, and ^DNISABA = *Halikiš* III 10, 21).

^DŠiwaz 'Day' occurs only in list 1. The god is at least indirectly represented by 'the gods of Day' in list 2 and 'the Heptad of the Day' in list 4 that have just been dealt with. That the god 'Day' really stands for 'Lucky Day' is indicated by the well-attested interchange between *šiwaz*, *šiwaz aššuš* and UD.SIG.³³ In X 81 10 f. we read ^DUD.SIG, *iš-pa-an-ta-an-na* 'Lucky Day and Night' instead. This, by the way, is an example for the veneration of the 'Night' as well, a fact significant for the appraisal of the 'Cappadocian' god *Išput*.

Šuwaliaz³⁴ is a minor deity connected with the house. This is clearest from the

and often contested since then, has finally prevailed. Bo 2372 (*JCS* 4.125) III 26 proves that, in spite of Ehelolf's remarks *OLZ* 1934.721 f., ^DKAL-aš-mi-iš (XX 24 IV 9) corresponds to ^DIn-na-ra-aš-mi-iš. Furthermore the sequence of gods in sacrificial texts shows that ^DKAL and ^DI-na-ar interchange (Laroche, *Rec.* 79 ff.; partly anticipated in F. Sommer, *AU* 21). The singer of Kaniš recites for ^DKAL also XXVIII 101 5 f. (Cf. further Otten, *Bibl. Or.* 8.230.)

³² KBo II 16 4. The milieu of this text is decidedly Luwian, note ^DTa-pár-ri-ia-aš-ši-iš (l. 5), ^DPt-ḫa-am-mi (l. 6) and the plural ^DU-ti-iu-ú-ni-en-zi (l. 8). There may also exist a Hurrian deity of this name (cf. von Brandenstein, *Bildbeschreibungen* 59 f.). *Maliya* is a frequent and presumably Hurrian name of males in the Nuzu texts (see Gelb, Purves, and MacRae, *Nuzi personal names* 95).

³³ Otten, *JCS* 4.121 ff.

³⁴ This is the nominative (correct Laroche, *Rech.* 60, accordingly), as it appears XX 65 3; *Šuwaliyatti* is dative. I do not see much reason to claim this god as Hurrian, as Laroche does. Before I can believe in the alleged identity of ^DNINURTA, I must see the unpublished text Bo 5841 from which the identification (von Brandenstein, *Bildbeschreibungen* 72 f.) is deduced. The data presented in the text make me rather sceptical.

various lists of 'holy places', a rather full form of which is read KUB X 11 IV 8 ff.³⁵

ZAG.GAR-RA-ni (i.e. *ištanani*) *piran* 'in front of the "altar"'

Dam-na-aš-ša-ra-aš 'to the D. (goddesses)',³⁶

^DŠu-wa-li-ia-at-ti 'to Š.',

KI.NE (i.e. *ḥašši*) *iš-tar-na pé-e-di* 'into the hearth',³⁷

^{GIŠ}TÁG-ti (i.e. *ḥalmašuitti*) 'to the throne',

^{GIŠ}AB-ia (i.e. *luttiya*) 'to the window',

^{GIŠ}PA DINGIR-LIM 'to the scepter of the god',³⁸

PA-NI ^{GIŠ}TUKUL (i.e. *ḥattalli?*) 'in front of the mace',

PA-NI ^{GIŠ}MAR.GÍD-DA 'in front of the wagon',

ḥa-at-tal-wa-aš GIŠ-i (i.e. *darui*) 'to the (door) bolt',

KI.NE (i.e. *ḥašši*) *ta-pu-uš-za* 'by the side of the hearth'.

A similar list is XXX 41 II 22 ff. (cf. IV 12 ff.):

^DŠu-wa-li-ia-at-ti 'to Š.',

[^DḤa-ša-m]i-li 'to Ḥ.',³⁹

KI.NE (i.e. *ḥašši*) [*piran*] 'in front of the hearth',

^{DUG}*iš-pa-an-tu-uz-zi* 'to the libation vessel',

ḥa-at-[tal-wa-aš GIŠ-i (i.e. *darui*) 'to the (door) bolt',

^{DUG}*ḥar-[ši-al-li]* 'to the ... vessel'

KI.NE (i.e. *ḥašši*) *ta-pu-uš-za* 'by the side of the hearth',

^DḤi-la-aš-ši 'to the demon of the ḥ.'⁴⁰

In XXXII 84⁴¹ I 14 ff., XXXIV 102 II 2 ff. and III 15 ff. and in IBoT I 23 III 6 ff. the enumeration begins with the lower ranks. The list runs, inverted, as follows:

^DUTU ^{URU}*Arinna*⁴² 'to the Sun-goddess of Arinna',

^DUTU (*ŠA-ME-E*) 'to the Sun-god (of heaven)',

^DEN.ZU⁴³ 'to the Moon-god',

^DÉ.A 'to the god of the Deep',

^DNISABA (= ^DḤal-ki⁴⁴) 'to the grain deity',

^DŠu-wa-li-ia-at-ti 'to Š.'

In XXX 41 cols. V and VI we encounter a series of gods which begins in what may be called the 'orthodox' fashion⁴⁵ and ends up with *Šuwaliyaz*, *Ḥašameliš*, the Storm-god and the Storm-god 'of the house'.

Just as ^DAššiyat-s belongs to the Hittite verb *aššiya-* 'love', ^DŠuwaliyat-s should belong to a verb **šuwaliya-*. To be sure, such a verb has not turned up

³⁵ KBo II 30 obv. 1 ff. runs parallel.

³⁶ See below.

³⁷ From here on also III 1 ff.

³⁸ The insertion III 13 ff. has been disregarded here.

³⁹ See below.

⁴⁰ See below.

⁴¹ Duplicate to XXXIV 102 IV.

⁴² This is broken off on XXXII 84 and IBoT I 123.

⁴³ IBoT I 23 III 17 offers DINGIR.MEŠ, which must be emended to ^DÉŠ, i.e. ^DXXX 'Moon'. In XXXII 84 either this or the following ^DEA is missing.

⁴⁴ Thus XXXIV 102 II 4; cf. HT 79 5.

⁴⁵ See Laroche, *Rec.* 80.

as yet; but a denominative of *šuwel* 'tie, string'—like *kururiya*- 'implement hostility' of *kurur* 'hostility'—would have to have this form. As we shall see presently, there is evidence that *Hašameliš*, with whom together⁴⁶ *Šuwaliyaz* appears in XXX 41 (see above), has to do with the door. Hence I would like to explain *Šuwaliyaz* as an analog of Akk. *ša uppi* 'the one of the uppu', i.e. a string device by which a door is opened.⁴⁷

The combination ^D*Šuwaliyatti* DINGIR.SAL.MEŠ-*ia* of Bo 2597 (list 1), with enclitic -*ia* meaning 'and', likewise fits in. These goddesses are possibly the *Damnaššaraš* of X 11 IV (see above) who, to judge by the formation of the word with the suffix -*šara*-, must be (at least originally) female and who seem to be attendants at the gate.⁴⁸

Hašam(m)iliš has already occupied us in connection with *Šuwaliyaz*. The most characteristic fact about this doubtless minor god is his ability, referred to in *Muršili's* annals,⁴⁹ to conceal armies on their march. His chthonic nature is apparent from his association with *Lelwaniš*, who is equated in Hittite texts with ^DEREŠKIGAL and with ^DALLATUM, both Mesopotamian goddesses of the nether world.⁵⁰ Among the lists in which *Hašamiliš* occurs, Bo 535/c etc. (JCS 4.123) is the fullest and offers *Lelwaniš*, [Lucky Day⁵¹], *Ašammaz*, *Tašammaz*, Sun-god, *Ištuštai*, *Pappai*, [*Hašameliš*], *Zilipuri*, *Šitarzuna*, etc. In the treaties he follows *Yarriš* and *Zappanaš* (*Zampanaš*, *Zapnaš*), of whom the first mentioned is the god of pestilence, i.e. a *Nerigal* figure. In 'orthodox' lists (see E. Laroche, *Recueil* 80) he is sometimes inserted between ^DX ^DX URU *Zippalanda* and ^DKAL = ^D*Inar(aš)*; thus XXV 6 III 22 (cf. XXXII 107 6).

The name *Hašamiliš* becomes more accessible to the lexicographer because the lists of 'holy places' occasionally mention the 'wood' (or wooden tool) of the god. See in particular XXV 18 III 28 ff., IV 24 ff., V 17 ff.:

iš-ta-na-ni 'to the "altar"',
ha-aš-ši-i 'to the hearth',
 KUB⁵² *kur-ši* 'to the shield',⁵³

⁴⁶ The combination of *Hašammiliš* with *Šuli(n)katte* in XXII 27 I 9 ff. (duplicate ABoT 14 II 8 ff.) seems accidental. The passage certainly must not be interpreted in the sense that the two deities are a divine couple. Instead of DUMU-RU-TI (Otten, *Bibl. Or.* 8.229 fn. 42) read É.MEŠ DINGIR.MEŠ-*ia* ku-e TUR-RU-TI (i.e. Akk. *šeḫḫerūti*) ŠA ^DŠ. Ū ŠA ^DH. 'what small temples there are of Š and H.'

⁴⁷ Sm 12 (CT XIX 23) obv. 5 ff. (names of professions):

<i>kak.š.lal</i>	= <i>na-du-ú</i> ⁴ <i>kakki</i>	'the one who lets the peg drop'
<i>kak.š.du₃</i>	= <i>pi-lu-ú</i> ⁴ <i>kakki</i>	'the one who opens the peg'
<i>kak.š.du₃</i>	= <i>mu-še-lu-ú</i> ⁴ <i>kakki</i>	'the one who raises the peg'
<i>lú.mud</i>	= <i>ša up-pi</i>	'the one of the uppu'
<i>lú.e</i>	= <i>ša nam-za-qi</i>	'the one of the key'
<i>lú.niḡ.kak.ti</i>	= <i>ša nam-za-qi</i>	'the one of the key'

⁴⁸ Occurrences: KBo II 30 obv. 1; KUB II 3 III 25; X 11 IV 10; XIV 8 I 35; XVI 34 I 6; XVII 1 II 13; XXIX 8 II 3 (here with determinative LÜ 'man'); XXX 42 I 19; XXXII 121 II 32 f.

⁴⁹ A. Götze, *Annalen des Muršiliš* 126 f. Note the apposition 'my lord'.

⁵⁰ Otten, JCS 4.119 ff. (1950).

⁵¹ Because of Bo 2593/c rev. 11 (l.c. 124).

⁵² See *ibid.* rev. 17.

⁵³ Or 'fleece'?

GIŠTĀG-ti (i.e. *ḫalmašuitti*) 'to the throne',
 GIŠAB-ia (i.e. *luttiya*) 'to the window',
^DḪa-ša-mi-li-ia-aš GIŠ-i (i.e. *darui*) 'to the "wood" of Ḫ.',
 TŪL-i⁵⁴ 'to the "spring"',
^{DUG}ḫar-ši-ia-li 'to the pithos',
 GIŠḫa-at-tal-wa-aš GIŠ-i (i.e. *darui*) 'to the (door) bolt',
 ḫa-aš-ši-i ta-pu-uš-za 'by the side of the hearth'.

Restore XX 39 1 ff. accordingly, perhaps also X 32 V 13 and 41 4 (in spite of the missing determinative).

In the circumstances, the assumption seems justified that the function of the god has to do with shutting out the light either of the sun or of the stars. The position of the 'wood of Ḫ.' after the window suggests that this may be accomplished by tightening the windows or their shutters. Note that Muršiliš says, when the hiding faculties of Ḫašamiliš are utilized, that the Storm-god summons Ḫašamiliš; it looks as though he may be in charge of the 'windows' of the heavenly palace which play such a prominent part in the palace of Ba' at Ugarit (there called 'urbi, hln).⁵⁵

In passing, it may be remarked that the name *Ḫašam(m)iliš* is curiously reminiscent of *Καδμῖλος*, *Κασμῖλος*, the name of the fourth of the *Καθεῖροι*—and servant of these 'great gods'—worshipped at Samothrace off the coast of Asia Minor.⁵⁶ Note the variant spelling ^DḪa-aš-mi-li-iš KBo I 2 rev. 25.

As was said before, *Ḫilaššiš* is clearly identical with Palaic *Ḫilanzipaš*. In one and the same group with *Kamrušipaš* and *Aškašipaš* one may rather expect the latter form. Instead, one reads *Ḫilaššiš* in Kaniš lists. The deity is certainly a minor deity and therefore often found at the end of enumerations. The name must be etymologically connected with the noun *ḫila-*, which probably denotes⁵⁷ the courtyard of a temple or palace.

The best preserved list of this kind is XXX 41 II 22 ff. quoted above. Compare also XX 99 III 1 ff.:

[^DŠu-wa-li-ia-at-ta-an]⁵⁸ 'Šuwaliyaz',
^DḪa-ša-mi-li-in 'Ḫašamiliš',
 [P.....] [.....],
 [^D.....⁵⁹] [.....],
^DḪal-ki-in 'the grain goddess',
 DINGIR.LŪ.MEŠ-eš 'the (male) gods',
 KINE (= *ḫaššan*) 'the hearth',
^DḪi-i-la-aš-ši-in, 'Ḫilaššiš'.

and a little more orthodox: XX 13 IV 3 ff.:

ḫa-aš-ši-i iš-tar-na pé-e-di 'into the hearth',

⁵⁴ Is this a cultic vessel—analogue to the 'Molten Sea' of Solomon's temple—or a specific individual spring?

⁵⁵ See most recently A. S. Kapelrud, *Baal in the Ras Shamra texts* 95 f. (1952).

⁵⁶ The similarity was noticed by Sayce, *JHS* 45.163 (1925); see also B. Hemberg, *Die Kabiren* 87, 316 f. (1950).

⁵⁷ Friedrich, *ZA* NF 3.179 ff.

⁵⁸ Cf. XXX 41 II 22 ^DŠu-wa-li-ia-at-ti, followed in the next line by [^DḪa-ša-mi-li-li].

⁵⁹ Quite likely ^DKAL is to be restored, cf. ^DKAL-ri pi[-ra-an] XXVIII 91 I 8 following ^DḪal-ki p[-ra-an] in l. 7. Note also the sequence ^DUD.SIG., ^DḪa-ša-am-me-li, ^DḪal-ki, ^DKAL of HT 79.

- A-N[A.....⁶⁰] 'to [.....]',
 GIŠAB-ia (i.e. *luttiya*) 'to the window',
 GIŠTĀG-ti (i.e. *ḫalmašuitti*) 'to the throne',
 DHi-la-[aš-ši] 'to Hilaššiš'.

In a group of '8' (l. 10) *Hilaššiš* appears in XXXII 87 rev. 11 ff.:

- D^{Ma-du-ša-ḫi-li-iš} 'Madušahiliš',
 D^{Da-a-ra-wa-[aš]} 'Darawaš',
 [.....]⁶¹ [.....],
 D^{Zi-li-pu-u-ri-iš} 'Zilipuriš',
 D^{EN.ZU M[UL]}⁶² 'the Moon (as a) star',
 [.....]⁶³ [.....],
 D^{Ha-ri-iš-ta-aš-ši-iš}⁶⁴ 'Haraštaššiš',
 D^{Hi-la-aš-ši-iš}⁶⁵ 'Hilaššiš'.

The appearance of *D/Tarawaš*⁶⁶ in this context should be noted.

⁶⁰ Perhaps *kur-ši* 'to the shield', cf. XXV 18 III (see above).

⁶¹ Perhaps [*Ḥašamiliš*], see Bo 2593/c (JCS 4.124).

⁶² Cf. KBo IV 13 VI (see presently); furthermore EN.ZU-an IX 38 I 5 and DXXX¹ IBoT I 23 III 17. Is *Ši-tar-zu-na*- Bo 535/c etc (JCS 4 123) also pertinent?

⁶³ Perhaps [*Gul-še-eš*], see IX 28 I 8, III 3 (see below).

⁶⁴ This is the clear nominative sing. of a formation in -š(š)iš which so frequently forms divine names (see Laroche, *Rech.* 68 ff.). Hence it seems necessary now to interpret D^{Ha-ri-iš-ta-ši(-in)-na} in the sequence A-NA D^{Gul-šu-uš} D^{MAḤ.MEŠ-uš} D^Ḥ. (KBo IV 13 II 20, IV 2 f., 22 f.) as an incorrect accusative in -in plus -a 'and' (cf. D^{Ha-ri-iš-ta-ši} after A-NA to be restored in III 6) and not as a Hurrian plural in -na (Goetze, *Tunnawi* 55 fn. 175). The noun on which *Harašta-š(š)i-* is based is unknown.

⁶⁵ The sequel will be dealt with below.

⁶⁶ This deity is also included in the list of KUB IX 28 which (I 9, III 6 f.) closes with 'the Thousand Hittite Gods'. It first enumerates some of the chief deities, namely:

- | | |
|-------------------------------------|---------------------------------|
| PUTU ^{URU} Arinna (I 3) | 'the Sun-goddess of Arinna', |
| P]UTU-uš (I 3) | 'the Sun-god', |
| D ^{X-aš} (I 3) | 'the Storm-god', |
| ne-pi-ša-aš [<i>tak.na-aš-ša</i>] | 'the gods of Heaven and Earth', |
| DINGIR.]MEŠ-mu-uš (I 3 f.) | |

and then proceeds to a special list:

- | | |
|--|-----------------------------|
| ne-pi-iš <i>te-e-kán</i> (I 4, II 26) | 'Heaven (and) Earth', |
| P]ISKUR-na-aš D ^{UTU} -uš (I 5, II 27) | 'Storm-god (and) Sun-god', |
| D ^{EN.ZU-an} (I 5) | 'Moon-god' |
| Ḥal-ki-iš (I 5) | 'Grain(-god)', |
| D ^{NIN.É.GAL} P ^{I-na-r} a-ša (I 5 f., II 28 f.) | 'Belat-ekallim and Inaraš'. |

-
- | | |
|---|----------------|
| D ^{IŠtar-iš} (I 7, III 1) | 'Ištar', |
| D ^{Nin-na-at-ta-aš} D ^{Ku-li-it-ta-aš} (I 7, III 2) | 'N. (and) K.', |
| D ^{Ta(-a)-ra(-u)-wa-aš} (I 8, III 3) | 'Tarawaš', |
| D ^{Gul-še-eš} (I 8, III 3) | 'the Gulšeš', |
| D ^{Hi-la-aš-ši-iš} (I 8, III 4) | 'Hilaššiš'. |
-

- | | |
|---|------------------------------|
| HUR.SAG.MEŠ (I 9, III 5) | 'The mountains', |
| ÍD.MEŠ (I 9, III 6) | 'the rivers', |
| ^{URU} D ^{Ha-at-tu-ša-aš} LI-IM DINGIR.MEŠ (I 9, III 6 f.) | 'the Thousand Hittite Gods'. |
-

The name of D^{Tarawaš} rather than that of the Hattic D^{Taru-} is probably contained in the 'Cappadocian' and Hittite proper name *Taruḫšuš* (< **tarawaḫšuš*).

A similar group of '15' (X 82 3) is found KBo IV 13 VI 31 ff. (restored by attaching X 82⁶⁷):⁶⁸

- ^DGul-šu-uš⁶⁹ 'The Gulšeš',
^DMAH.MEŠ-uš⁷⁰ 'the MAH goddesses',
^DKu-za-na-šu 'Kuzanašu',
^DU.GUR Ū ^DU.GUR URU⁷¹Ha-ia-ša 'U.GUR and U.GUR of Hayaša',
^DXXX MUL 'the Moon (as a) star',
iš-pa-an-za 'the Night',
^DHa-aš-ma-i-ú-un 'Hašmayuš',
^DKa-taḥ-ḫi 'Kataḫḫi',
^DHa-ri-iš-la-aš-ši-in 'Harištaššiš',
^DḪi-la-aš-ši-in⁷¹ 'Ḫilaššiš'.

The reading of ^DU.GUR (list 1) in Hittite, Luwian, Palaic, or Kanishite is unknown;^{72a} in Hattic, we are told by some scholars (F. Sommer, *OLZ* 1939.685; E. Laroche, *Recherches* 31), it should be *Šulinkatte*. However, this becomes doubtful when one compares the sequences:

- ^DX ^DWa-a-ši-iz-za-li-in ^DU.GUR ^DWa-a-ḫi-ši-in II 3 II 38 f.
^DU.GUR ^DWa-a-ḫi-ši-in XI 35 V 23

[^DX] ^DWa-a-[ši-iz-za-li-in] ^DWa-a-a[?] ^DWa-a-ḫi-ši-in IBoT II 77
(continuation of X 67 V). One might think of using ^DWa-š[u-um-ma-an'] (XX 48 VI 14) for restoration.

The Kanishite(?) reading may be indicated by the sequence

[^DHa-š]a-am-me-li, [^DḪi-la-aš-ši, [^DHa-pa-an]-tal-li-ia-a II 13 III 1 ff.
when compared with *ibid.* III 26 ff. (list 2 above). *Ḫapantalliyaš*, to judge by XXVIII 3-5, belongs indeed—as a personal attendant?—to *Kataḫziburi* = *Kamrušepaš*.⁷²

U.GUR as a Sumerogram is supposed—see the vocabulary Sb 210—to equal Akk. *namšaru* 'sword'. Hittitologists can hardly help thinking of the 'swordgod' pictured in the side gallery at Yazilikaya.

A number of other passages in which the 'singer of Kaniš' officiates center around a river or a mountain, and offer other points of contact with the lists that have just been discussed. These passages, never treated satisfactorily before, will here be presented in full.

⁶⁷ See Bossert, *Königssiegel* 33.

⁶⁸ Most members of this list are in the accusative.

⁶⁹ Cf. Goetze, *Tunnawi* 62 f.; Friedrich, *JCS* 1.280 ff.; Laroche, *Rech.* 98 f. and *JCS* 2.125. For the reading, Palaic *Gul-za-an-ni-ke-eš* (Bo 1551/c etc. [ZA NF 14.139 ff.] rev. 18) is of importance. It shows that 'gul' is to be read thus, since complements of such length are unusual.

⁷⁰ It is very unlikely that these goddesses bear the same name as ^DMAH = *Hannaḫannaš* (Otten, *OLZ* 1941.18; Laroche, *Rech.* 101). For the reading, note ^DU^DUTU^DKAL ^DGul-šu-uš [^Dx-š]a-ra-uš-ša VII 2 I 15 f.

⁷¹ For the sequel see below.

^{72a} U.GUR may be of Hurrian origin. Cf. Gelb-Purves-Macrae, *Nuzi personal names* 271.

⁷² *Ḫa-pa-li* occurs as a proper name in the Kültepe texts: TCL I 240 1 ff.

IBoT II 19:

- [x x x -h]a-aš ^DHu⁷³-wa-ri-ia-an[-zi-pa-an⁷⁴]⁷⁵
 2 [G¹⁵KĀ.G]AL-aš ^DŠa-li-wa-ni-iš [.]
 [^DNa-ak-ki-li-ia ^ÍD-aš ^D[KAL]
 4 [DINGIR.LÚ].MEŠ ŠA ^DÉ.A KI.NE ^DG[ul-še-eš]⁷⁴
^DHi-i-li¹-aš-ši-in ir-ḥa-a-du[-uš]
 6 [ḥe-eš]-ša-al-la-an-du-uš⁷⁶ ^D[UD.SIG₆-ia]
 [.] 1-ŠU a-ku-an-zi [.]
 8 [LÚNAR] ^{URU}Ka-ni-[iš SĪR-RU]

Ḥuwariyan[zipaš] of [.]ha,

- 2 The Š. deities of the gates,⁷⁶
 the river Nakkiliya, [Inar] of the River,
 4 the (male) gods of EA,⁷⁷ the hearth, the G[ulšeš],
 Ḥilaššiš, those included
 6 (and those) excluded,⁷⁸ [and the Lucky Day]
 [they] "drink" once;
 8 [the singer of] Kani[š sings]."

KUB IX 21 (// ABoT 3⁷⁹):

- [EGIR-ŠU-]ma ^ÍD-an ^ÍD-aš ^DKAL ŠA ^Í[D⁸⁰]
 4 [DINGIR.MEŠ ḥ]a-an-ti-ia-aš-ša-aš ^DÚR-aš 3-ŠU [e-ku-zi]
 [3^{NINDA}ṭ]a-kar-mu-uš pár-ši-ia ^{LÚ}N[AR ^{URU}Ka-ni-eš SĪR-RU]
 6 [EGIR-ŠU-]ma wa-ap-pu-u-wa-aš ^DGul-še[-eš]
 [^DÚR-aš 3-]ŠU e-ku-zi 3^{NINDA}ṭa-k[ar-mu-uš pár-ši-ia]
 8 [^{LÚ}NA]R ^{URU}Ka-ni-eš SĪ[R-RU]
 [EGIR-ŠU-m]a ^ÍD (^ÍD-a[š ^DKAL ^ÍD]-aš⁷¹)
 10 [(DINGIR.M)EŠ ḥa-an-ṭi-ia-(aš-ša-aš ir-ḥa-an-d)u-uš
 [ḥe-eš-ša-al-la-an-(du-uš 3-ŠU e)-ku-zi]

⁷³ The restoration is uncertain; cf. ^DHu-u-ri-ia-an-zi-pa-aš V 7 obv. 17; ^DHu-ri-ia-an-zi-pa-an XX 4 I 13; ^DHu-ri-ia-an-zi-pt ibid. 16.

⁷⁴ The restoration is uncertain.

⁷⁵ Cf. II 8 V 29.

⁷⁶ G¹⁵KĀ.GAL DINGIR.MEŠ ṣa¹-la-wa-ni-iš (plur. nom.) II 8 V 24; KĀ.GAL-aš ^Dṣa¹-la-u-wa-ni-iš (plur. nom., as voc.) X 89 I 15; G¹⁵KĀ(.GAL)-aš DINGIR.MEŠ ṣa-la-wa-ni-ia-aš (plur. dat.) II 8 I 25, II 43, cf. XXV 49 III 25.

⁷⁷ See above.

⁷⁸ For ḥešalla- cf. KBo IV 14 III 13 ff. '[if] something troublesome happens to the king, either that the king falls ill, or that the lands become rebellious, or that the enemy invades the land, then the blocked [ways] (KAS.ḤĀ-ta-ma ḥi-iš-šal-al-la) let in the same way be your concern.' Restoration and interpretation are supported by XXXI 136 rev. 4 ff.: 'Should something troublesome happen to the Labarna, the k[ing, (either that) the ways of the earth are blocked to you (tu-uk ḥe-e-ša-la⁷²-ṭi⁷²-iš] (or) the ways of [.] are left open(?) (tar-na-ti-iš)'

⁷⁹ ABoT 3 3 = KUB IX 21 9; words preserved on ABoT 3 are between parentheses. For KUB IX 21 alone cf. Bossert, *Heth. Königssiegel* 43. A similar text is KUB X 50.

⁸⁰ Compare I. 9.

12 [³ NINDA *ta-kar-mu-uš* (*pár-ši-ia* LÚNA)R ^{URU} *Ka-ni-eš* SÌR-RU]

'Afterward he/she "drinks" the River,⁸¹ the Inar of the River (and of the R[iver]

4 the male(?)⁸² [gods] in a sitting position three times; he breaks three *t*-[loaves]; the s[inger of Kaneš sings].

6 [Afterward] he "drinks" the Gulšeš of the clay pit
[in a sitting position three] times; he breaks three *t*-loaves

8 [the sin]ger of Kaneš si[ngs].

[Afterward] he "drinks" the River, [Inar] of the River (and) of the River]

10 [the m]ale [gods], those to be included
[and those to be exclud]ed, three times;

12 he breaks [three *t*-loaves;] the sing[er of Kaneš sings].'

Compare also IBoT I 2 III 10 ff.⁸³

10 LUGAL-*uš* ³-*e*

ir-ḫa-a-u-wa-an-zi

12 ^{1D} *Ma-ra-aš-ša-an-da*

^DKAL ^{1D} *ka-lu-ti*

10 'The king (proceeds)
to treat three:

12 the Maraššandaš River,
Inar of the River, and (their) circle⁸⁴.'

'River' in IX 21 probably means the Maraššandaš.

In a wider sense II 8 I 20 ff.⁸⁵ belongs here, at least the second part of this list which centers around the River Zuliyaš. The whole list, however, may find its place here:⁸⁶

A-NA ^DKAL ^{URU} *Ta-ú-ri-iš* 'to Inar of Tauriš',

[A-NA AMA K] *a-li-im-ma*⁸⁷ 'to mother K.',

[^DḪa-ša-me-li]⁸⁸ 'to Ḫašameliš',

^DA-*aš-ši-ia-za*⁸⁹ 'Aššiyats',

⁸¹ Probably the Maraššandaš river is meant.

⁸² Cf. ŠA *La-ba[-ar]-na ḫa-an-te-ia-aš-ša-aš-ši-iš* ^DKAL-*i* KUB II 1 II 47. The formations in -šiš, structurally adjectives but syntactically equivalent to Hitt. genitives, like *iš-ḫa-šar-wa-na-aš* (see KBo II 38 III 9) in the preceding line, appear in this text in great numbers and qualify the often repeated dative ^DKAL-*i*; curiously enough they are never inflected themselves, and apparently remain in the nominative.

⁸³ See also Bossert, *Heth. Königssiegel* 48. In IV 1 the river is mentioned by name again, but in IV 11—with the singer of Kaneš singing—he is merely referred to as 'the River'.

⁸⁴ See Goetze, *JCS* 1.87 f.; Laroche, *JCS* 2.113.

⁸⁵ Easily restored from *ibid.* II 37 ff., V 13 ff.

⁸⁶ See also Bossert, *Heth. Königssiegel* 33 ff.

⁸⁷ Probably also XX 48 I 3 (Bossert, *l.c.* 45, where l. 5 must begin with *ku-it-ma-an-ma* 'as long as!'); in l. 7 the singer of Kaniš sings.

⁸⁸ See V 20.

⁸⁹ This nominative is rigidly maintained as though it were an ideogram. In II 40 one even reads A-NA ^DA-*aš-ši-ia-za*, and in V 22 ^DA-*aš-ši-ia-za* serves as an accusative. In the latter place, strangely enough, the items that follow, syntactically accusative, also appear in the nominative.

^DX ^{GIŠ}TIR 'the Storm-god of the Grove',⁹⁰

[^{ID}Zu-]i-ia ^DKAL ^{ÍD}-ia 'River Z. and Inar of the river',

^{GIŠ}KÁ.GAL-aš ^{DINGIR.MEŠ}ša-la-wa-ni-ia-aš 'to the š. deities of the gate',

[te-pu pé]-e-da-an 'the small place',⁹¹

EME-aš ha-an-ta-a-an-za 'true "tongue"',

an-na-ri-iš tar-pi-iš

zi-pu šar-ru-mar 'dividing the z.',

^DUTU.SIG₅-ia 'and Lucky Day'.

The name of *Aššiyats* is certainly reflected by the Kültepe name *A-ši-i-el* etc.

With a mountain we have in similar fashion XX 48 VI 10 ff.⁹²

10 LUGAL SAL.LUGAL DÚR-aš 6 a-ku-wa-an-zi

^{HUR.SAG}Ta-pa-la ŠA ^{HUR.SAG}Ta-pa[-la]

12 ^DKAL DINGIR.LÚ.MEŠ ^DMa'-li'-i[a]

ŠA DINGIR.LÚ.MEŠ ^DÉ'[A...]⁹³...

14 ^DWa-šu-um-ma[-an? ^D.....]

10 'The king and the queen "drink" in a sitting position six:
the mountain Tapala,⁹⁴ Inar of the mountain Tapala,

12 the (male) gods of Maliyaš(?),
the (male) gods of (?) E[a.]

14 Wašumma[š (and)]

And finally there is with ^DKAL ^{KUŠ}kuršaš in a remotely similar enumeration IBoT II 22 4 ff.:

4 [nu LU]GAL-uš GUB-aš GAL-it ^DK[AL ^{KUŠ}kur-ša-aš⁹⁵]

[ir-ḫ]a-u-wa-a-ar ir-ḫa-an-du-uš [DINGIR.MEŠ-us]

6 [e-ku]-zi LÚ.MEŠNAR ^{URU}Ka-ni-i[š ŠĪR-RU]

[NINDA₅]a-kar-mu-un pár-ši-ia na-an [.]

8 [A-NA PA-]NI DINGIR-LIM da-a-i

4 'The king "drinks" in a standing position with a cup Inar [of the shield]
(and) [the gods] due for treatment;

⁹⁰ The festival in question takes place in the 'Grove of Tauriš' (VI 4).

⁹¹ This and the three following items remain obscure, although *tepu pedan*, EME-aš *hantanza*, and *zipu šarrumar* recur in KBo IV 13 (+ X 82) VI 36 f. and all four in X 81 6 ff. (where l. 8 *an-na-ri-iš* with a preceding 'Glossenkeil' is noteworthy). It is of course no accident that in the long list of KAL deities of the Labarna (KUB II 1) we find:

ŠA La-ba-ar-na ^DKAL te-pa-u-wa-aš pé-e(-da)-aš (II 39 f.),

ŠA La-ba-ar-na ha-an-da-at-ta-aš ^DKAL-i (II 45),

ŠA La-ba-ar-na ha-an-te-ia-aš-ša-aš-ši-iš ^DKAL-i (II 47),

ŠA La-ba-ar-na tar-pa-at-ta-aš-ši-iš ^DKAL-i (II 48).

Cf. also ŠA La-ba-ar-na-aš [x x x -]aš tar-pa-aš ^DKAL (III 1 f.), which might have read [an-na-ra-]aš. For *an-na-ri-in tar-pi-in*, again in a god list, see also XXXII 87 rev. 14. One can hardly avoid being reminded of the ^DAn-na-ru-um-mi-en-zi, Luwian for ^DInnarawanteš of IX 31.

⁹² See also Bossert, *Heth. Königssiegel* 45. At the beginning of col. VI there precede ^DHé-bat (Hurrian singer) and ^DKAL SE-RI 'Inar of the Field' (singer of Kaniš).

⁹³ Uncertain, cf. ŠA DINGIR.MEŠ LÚ.MEŠ ^DMa-li-ia (dat.) KBo IV 13 I 16.

⁹⁴ Also XXXV 32 I 19.

⁹⁵ Restored according to l. 3.

- 6 the singer of Kaniš sings;
 he breaks a *t.*-loaf and places it [on the table]
 8 [in fr]ont of the deity'

The text KUB XX 60⁹⁶ has gained in importance since it has been fitted to IBoT I 22 and IBoT II 20. It now reads:

- 8 [LUGAL-uš GUB'-aš ^D]KAL e-ku-zi
 [L^UNAR ^{URU}K]a-ni-iš SİR-RU
 10 [1 NINDA.KUR₄-RA p_{ár}-ši-i]a' NINDA^azi-ip-pu-la[-aš-ši-in]
 [NINDA] NINDA^ata-p_{ár}-wa₆-šu-un-n[a]
 12 [A-NA GIŠBANŠUR ti-a]n-zi
 LUGAL-uš GUB[-aš ^DX] ^{URU}Ne-ri-ik[e-ku-zi]
 14 L^UKALA SİR-RU 1 NINDA.KUR₄-RA p_{ár}-ši-i[a]
 NINDA^ata-p_{ár}-wa₆-šu-un ti-ia-an-zi
 16 LUGAL-uš GUB-aš ^DIŠTAR ^{URU}Ša-m[u-ša e-ku-zi]
 L^UNAR ^{URU}Hur-ri SİR-RU 1 NINDA.KUR₄-R[A p_{ár}-ši-ia]
 18 LUGAL-uš GUB-aš ^DX KARAŠ e-ku[-zi]
 GIŠ.^DINANNA GAL SİR-RU 1 NINDA.KUR₄-RA p_{ár}-ši-ia
 20 LUGAL-uš GUB-aš 7 ^DX.ĤÁ ^DX[.]
^DX ŠA-ME-E ^DX mu-wa-tal-la[-hi-ta-aš]
 22 ^DX ku-un-na-aḥ-ḥu-u-wa[-aš] ^DX x[.]
^DX ḥa-an-da-an-da-a[n-n]a-aš
 24 ^DX wa-ar-ra-ḥi-ta-aš e[-ku-zi]
 GIŠ.^DINANNA GAL SİR-RU 1 NINDA[.KUR₄-RA p_{ár}-ši-ia]
 26 LUGAL-uš GUB-aš ^DIB e-ku-zi
 GIŠ.^DINANNA GAL SİR-RU 1 NINDA[.KUR₄-RA p_{ár}-ši-ia]
 28 LUGAL-uš DŪR-aš DINGIR.MEŠ LŪ.MEŠ[. e-ku-zi]
 L^UNAR ^{URU}Ka-ni-iš [SİR-RU]
 30 1 NINDA.KUR₄-RA p_{ár}-ši-ia

A translation seems hardly necessary. The sequence of gods includes (l. 8 ff.) Inar(aš)⁹⁷ and then a series of seven Storm-gods (l. 20 ff.), the god IB⁹⁸ (l. 26 f.), and a group of '(male) gods' (l. 28 ff.), possibly those 'of EA' (see above). The fact must be stressed that some of the names of the Storm-gods are in Hittite and others in a language that one would label Luwian; it might be Kanishite, provided the assumption can be made that the notation 'singer of Kaniš' (l. 29) is valid for the preceding section also. The names in question are:

^DX mu-wa-tal-la[-hi-ta-aš] 'the St. of fierceness',⁹⁹

^DX ku-un-na-aḥ-ḥu-u-wa-aš 'the St. of guiding right',

⁹⁶ See also Bossert, *Heth. Königssiegel* 46.

⁹⁷ The singer of Kaniš sings for ^DKAL also XXVIII 101 5 f.

⁹⁸ Cf. von Brandenstein, *Bildbeschreibungen* 72 f. with fn. 7. The text Bo 3533 which he quotes is still unpublished.

⁹⁹ Cf. ŠA La[-b]a-a[r-]na mu-wa-ad-da-la-ḥi-da-aš ^DKAL-i II 1 III 14 f. As for the translation of *muwatalla/i* = Akk. *mutallu* (= NIR.GÁL), because of Bossert, *Heth. Königssiegel* 73, the fact must be recalled that Akk. *etellu*, and the *mutallu* derived from it, have nothing

⁹⁹ *PX* *ha-an-da-an-da-an-na-aš* 'the St. of glory',¹⁰⁰

⁹⁹ *PX* *wa-ar-ra-ḫi-ta-aš* 'the St. of help'.¹⁰¹

It appears, then, that the main gods of Kaniš are *Kamrušipaš*, *Pirwaš*, *Aškašipaš*, *Maliyaš*, and the various Heptades. *Pirwaš* is certainly identical with the 'Cappadocian' *Perwa*. We have furthermore encountered among them *Inar(aš)* and the pair *Šiwat* '(Lucky) Day' and *Išpant-* 'Night'. In related material we found *Aššiyaz*, *Ḫalkiḫ*, and *Tarawaš*. It is particularly worthy of note that *šiwat-* 'day', *išpant-* 'night', and *ḫalki-* 'grain' form part of the Hittite vocabulary, and that at least the first two exhibit Indo-European structure and have very good Indo-European etymologies. The names *Kamrušipas* and *Aškašipaš* are analyzable; they contain word stems otherwise known from Hittite and a suffix which is sometimes added to words of Indo-European structure and has been claimed above for Kanishite.

Out of the list of gods with which we started, then, only *Ilaliya* has not turned up in Kanishite surroundings. This is probably only due to the limited material at our disposal. *Ilaliya* is certainly reflected by the plurals Hittite ⁹⁹*Ilaliyanteš* and Palaic ⁹⁹*Ilaliyantikeš*. The goddesses¹⁰² of this name always appear together with the Sun-god, and seem to be servants of his.¹⁰³ Their name cannot be separated from the verb meaning 'desire, wish' which is known to us from Hittite as *ilaliya*.¹⁰⁴

Thus it can be stated that the theophorous elements that form part of the names of Anatolians occurring in Cappadocian tablets are Kanishite in nature. This is an eminently reasonable result, for the Assyrian trading center lies under the very walls of Kaniš. The linguistic appearance of these elements leaves no doubt that (Proto-)Indo-Europeans were present in Cappadocia as early as the period of Assyrian colonization, i.e. in the 20th century B.C.¹⁰⁵

to do with *eḫlu* '(sexually) mature, manly'. The singer of Kaniš sings for *PX* *NIR.GÁL* XXV 41 V 13 f.

¹⁰⁰ In the theological sense; see also [*ŠA* *La-ba-ar-n*]a *pa-ra-a ḫa-an-da-an-da-an-na-aš* *PKAL-i* II 1 II 20 and [*ŠA* *La-ba-ar-na* [*PA*-a-la-aš *pa-r*]a-a *ḫa-an-ta-an-ta-an-na-aš* *ibid.* IV 6 f.

¹⁰¹ This is clearly the reading for *PX* *RE-ŠU-TI* or *Á.DAH* of the treaties. This whole series recalls in a striking manner the series attributed to El in the Ugaritic text, *Syria* 14.231 ff. = 20.129 ff. (C. H. Gordon, *Ugar. handbook* No. 107): *b-mrḫ il*, *b-nit il*, *b-šmd il*, *b-dḫn il*, *b-šrp il*, *b-knt il*, *b-ḡdyn il*. They have found very divergent interpretations; see E. Dhorme, *Syria* 14 l.c.; J. Obermann, *JBL* 55.21 ff.; C. H. Gordon, *Ugar. literature* 109; O. Eissfeldt, *El im ugar. Pantheon* 60 f.

¹⁰² Otten (*ZA* NF 14.128 fn. 19) contended that *-k(a)*, which occurs in the plurals *Ilaliyantikeš*, *Gulzannikeš* and *Uli(li)yantikeš*, is the plural suffix of Palaic. I think he is mistaken. The second of these plurals is certainly feminine, and the 'Cappadocian' names in *-l(i)ka* and *-annika*, all feminines, suggest that *-ka* marks the feminine gender.

¹⁰³ *DI-la-li-ia-an-du-uš* (plur. acc.) IX 34 III 35; *DI-la-li-ia-an-ta-aš* (plur. gen.) II 4 V 27 (Palaic ritual!).

¹⁰⁴ Sommer, *Bogh. Stud.* 7.55.

¹⁰⁵ This article reaffirms a view which I expressed as early as 1931 and have intermittently defended ever since; see *ZA* NF 6.260 ff. (1931); *Kleinasien* 55, 68 f., 167 (1933); *Haverford symposium* 147 (1938); *Proc. of the Amer. Philos. Soc.* 97.000 (1953). Similar views are held by S. Alp, *Belleten* 50.269 f. (1949) = *JKF* 1.135 f. (1951), and by Bilgiç, *AfO* 15.18 f. (1951). When I received Alp's review in *Ankara Üniversitesi Dil ve Tarih-Coğrafya Fakültesi Dergisi* 10.249 ff. (1952), the present article had been long since completed.

IMPERFECT AND PRETERIT IN TOCHARIAN

GEORGE S. LANE

University of North Carolina

[The Proto-Tocharian imperfect, based on IE optative formations, is preserved in Kuchean (Toch. B) but lost in Turfanian (Toch. A) except for two relics, the imperfects of the verbs 'be' and 'go'. The remaining Turfanian imperfects are in origin identical with preterit formations found in Turfanian or in Kuchean or in both, which are all derived from IE perfects and aorists.]

As I have stated several times on other occasions, the more one attempts to compare the two Tocharian 'dialects' with a view to concluding something about the nature of Proto-Tocharian and its position among the Indo-European languages, the more one becomes impressed with the remarkable divergence of development which took place from the period of Proto-Tocharian unity to the time between 500 and 800 A.D. when the dialects appear. About all that really connects the language of Turfan in the east to that of Kucha in the west is a close similarity of grammatical system and a fairly high coincidence of vocabulary. Everywhere we find either the same grammatical function served by morphemes of entirely different phonetic shapes and of different formal origins, or else formal elements the same in origin but serving different grammatical functions. Nowhere is this situation more apparent than in the verbal system.

Both Turfanian (Toch. A) and Kuchean (Toch. B) possess four modes—indicative, subjunctive-future, optative, and imperative—and, in the indicative, a present, an imperfect, and a preterit tense. But there the coincidence all but ceases. There are of course some close parallelisms in stem formations in most verbal categories, but there are also marked discrepancies; and of all the verbal categories, it is the formation of the imperfect in Turfanian and its relation to the present system on the one hand and to the preterit system on the other that is most confounding. It is also in the formation of the imperfect that the two languages go farthest asunder. The present paper treats the identification of so-called imperfects and preterits and their ultimate origins. Some of the facts are obvious and already pretty well agreed upon, others are in dispute, and some of the problems have received no attention at all. In our treatment we shall in general proceed from what is most certain to what is more obscure.

The formation of the imperfect in Kuchean seems clear enough. It is formed by the addition of a characteristic sign *-i-* (*-ī-*, *-y-*) or *-ai-* (*-ey-*) to the present stem. Before this sign, an immediately preceding consonant or consonant group is subject to palatalization: *k* > *ś*, *nk* > *ñc*, *tk* > *cc*, *sk* > *śś*, *t* > *c*, *tl* > *cc*, *nt* > *ñc*, *n* > *ñ*, *l* > *ly*, *s* > *ś*.¹ Exceptions to this rule are the imperfects of Present Classes III and IV (cf. below),² and the present stems in original *ā*, which form the im-

¹ For further details of the palatalization, cf. *Lg.* 21.81 ff.

² For the convenience of the reader who may want additional information not pertinent to this paper, the numbering of classes (whether present, preterit, subjunctive, or optative) follows the system advanced by W. Krause, *Westtocharische Grammatik*, Band 1. *Das Verbum* (Heidelberg, 1952; abbr. Krause).

perfect in *-oy-*. A brief survey will elucidate the formation: Cl. I *kālñ-* 'resound', pres. pl. 3 *kalñem*, impf. act. sg. 3 *kalñi*; *klāñk-* 'doubt', pres. mid. sg. 3 *klyenträ*, impf. act. sg. 3 *klyeñci*; *pālk-* 'shine', pres. act. sg. 3 *palkām*, impf. act. sg. 3 *palyši*; Cl. II *aik-* 'know', pres. mid. sg. 1 *aikemar*, impf. *aışmar*; *klyaus-* 'hear', pres. act. sg. 3 *klyaušām*, impf. sg. 1 *klyaušim*; Cl. III *sruk-* 'die', pres. mid. sg. 3 *sruketrä*, impf. mid. pl. 3 *sruky(enträ)*; *māšk-* 'be', pres. mid. sg. 3 *māskētär*, impf. mid. sg. 3 *māskūtär*; Cl. IV *kārp-* 'descend', pres. mid. sg. 3 *korpotär*, impf. mid. pl. 3 *korpyentär*; *yāt-* 'be capable', pres. mid. sg. 3 *yototär*, impf. mid. sg. 3 *yotiträ*; Cl. V *kwā-* 'call', pres. mid. sg. 3 *kwātär*, impf. mid. sg. 3 *kwoytär*; *šwā-* 'eat', pres. act. sg. 3 *štwañ*, impf. act. sg. 3 *šwvoy*; Cl. VI *kārs-* 'know', pres. act. sg. 3 *kārsanam*, impf. act. pl. 3 *kārsanoyem*; *tärk-* 'release', pres. act. sg. 3 *tärkanam*, impf. act. sg. 3 *tärkänoy*; Cl. VII *pärs-* 'sprinkle', pres. act. sg. 3 *prantsām*, impf. mid. sg. 3 *prantsitär*; Cl. VIII *er-* 'produce', pres. act. sg. 3 *eršām*, impf. mid. pl. 3 *ersyenträ*; *lik-* 'wash', pres. mid. sg. 3 *lyikštär*, impf. mid. sg. 3 *likštär* (caus. ?); Cl. IX (primary) *ai-* 'give', pres. act. sg. 1 *aiskau*, 3 *aışšām*, impf. act. sg. 3 *aışši*; *yām-* 'do', pres. act. sg. 1 *yamaskau*, 3 *yamaššām*, impf. act. sg. 3 *yamašši*; *kālp-* 'get', pres. act. sg. 1 *kālpāsk(au)*, sg. 3 *kālpāššām*, impf. act. sg. 2 *kālpāššit*, 3 *kālpāšši*; (causative) *kārs-* 'know', pres. act. sg. 3 *šarāššām*, pl. 3 *šarāškem*, impf. act. sg. 3 *šarāšši*; *wātk-* 'decide', caus. 'command', pres. act. sg. 1 *watkāskau*, 3 *watkāššām*, impf. act. sg. 3 *watkāšši*; Cl. X *lāt-, lānt-* 'go out', pres. act. sg. 1 *lnaskau*, 3 *lnaššām*, impf. act. sg. 3 *lnašši*; *māl-* 'oppress', pres. mid. sg. 3 *mällästrä* (*ll < ln*), impf. mid. sg. 3 *mälläššitär*; Cl. XI *āks-* 'announce', pres. act. sg. 1 *aksaskau*, 3 *aksaššām*, impf. act. sg. 3 *aksašši*; Cl. XII *mānt-* 'injure', pres. act. sg. 3 *māntam*, pl. 3 *māntaññem*, impf. mid. sg. 3 *māntaññiträ*.

To these 'regular' imperfects are to be added those of the verbs 'be' and 'go':³

	'be'		'go'	
	Sg.	Pl.	Sg.	Pl.
1	<i>šaim, šeym</i>	<i>šeyem</i>	<i>yaim</i>	
2	<i>šait</i>	<i>šaicer, šeycer</i>	<i>yait</i>	
3	<i>šai, šey</i>	<i>šeyem, šem</i>	<i>yai, yey</i>	<i>yeyem</i>

The corresponding imperfects are found also in Turfanian, where these two verbs are the only examples of the 'i-imperfect':⁴

1	<i>šem</i>	<i>šemäs</i>	<i>ye(m)</i>	
2	<i>šet</i>		<i>yet</i>	
3	<i>šeš</i>	<i>šeñc</i>	<i>yeš</i>	<i>yeñc</i>

It is generally assumed that the origin of this imperfect is to be sought in an optative, showing either an extension of the Indo-European 'weak' athematic optative sign *ī* or possibly the thematic *oi*, reduced to *ī* in atonic position, or a combination of both.⁵ There is no reason to doubt this on either formal or semantic grounds. Pedersen has pointed out the similarity with the iterative use of the optative in Greek in temporal and relative clauses, and Couvreur has

³ Krause 255 and 222 respectively.

⁴ Cf. Sieg, Sieglings, and Schulze, *Tocharische Grammatik* (abbr. SSS) 384 ff.

⁵ Cf. Krause 103-4, 113-4.

indicated certain Indo-Iranian parallels.⁶ Krause has recently noted also the coincidence of the Tocharian development with that which may be assumed for certain of the British Celtic imperfects.⁷ In order to add greater plausibility to this view (if it is still doubted on semantic grounds), I would call attention to the use of the English conditional with *would* as a sort of habitual past, as in *When I was on vacation I would take a walk before breakfast every day, When we were in New York last month we would see a show every night.*

It is to be assumed, then, that this imperfect is a reflex of the Indo-European optative (thematic or athematic or both, cf. above). Its endings, however, except the 1st and 3rd sg. in Kuchean, are identical with those of the present indicative, e.g. from *camp*- 'be able' (act.), *aik*- 'know' (mid.):⁸

ACTIVE			MIDDLE	
	PRESENT	IMPERFECT	PRESENT	IMPERFECT
Sg. 1	<i>campau</i>	<i>cāmpim</i>	<i>aikemar</i>	<i>aišimar</i>
2	<i>campāt</i>	<i>campit</i>	<i>aištar</i>	<i>aišilar</i>
3	<i>campāņ</i>	<i>campi</i>	<i>aištār</i>	<i>aišilār</i>
Pl. 1	<i>campem</i>	<i>campim</i>	<i>aikemt(t)ār</i>	<i>aišyemt(t)ār</i>
2	<i>campcer</i>	<i>campicer</i>	<i>aištār</i>	<i>aišilār</i>
3	<i>campem</i>	<i>cāmpyem</i>	<i>aikentār</i>	<i>aišyentār</i>

It is noteworthy that *campim* and *campi*, the two forms of the imperfect which contrast with the indicative, also show what may be readily interpreted as the (originally athematic) primary ending *-mi* and the secondary (thematic and athematic) ending *-t*. The latter is of course proper to the optative; the former shows, as in Greek (Attic and elsewhere), the transfer from the μ inflection (e.g. $\phi\acute{\epsilon}\rho\omicron\mu\iota$ as opposed to the historically correct Arcad. $\acute{\epsilon}\xi\epsilon\lambda\acute{\alpha}\nu\omicron\iota\alpha$).⁹

It is debatable whether this imperfect use of the optative was a feature of Proto-Tocharian lost in Turfanian except for the roots *s*- and *i*-, or whether it was a special development in Kuchean after the separation of the dialects, originating in these two roots. I should think, however, and I believe the arguments of this paper will show, that the former is the more reasonable view. This would seem a logical conclusion, since in many other respects Kuchean is the more archaic of the two languages.

This being my assumption, I am compelled to look for the origin of the so-called imperfect in Turfanian, which must then be secondary. Perhaps a review of these formations will be helpful; the following outline retains the order and numbering of SSS:¹⁰

I includes the imperfect of *s*- 'be' and *i*- 'go', cf. above.

IIa lengthens the radical vowel and palatalizes the initial consonant if possible, and adds the stem vowel *a* before the endings (no ending in act. sg. 3): act. sg. 3

⁶ Pedersen, *Tocharisch* 204 f.; W. Couvreur, *BSL* 39.247 f.

⁷ Krause, *Journal of Celtic studies* 1.24 ff. Cf. also Pedersen, *VKG* 2.348.

⁸ Several of these forms are of course assumed for the particular verbs for the sake of uniformity of comparison.

⁹ Cf. Buck, *Introduction to the study of the Greek dialects*² 112.

¹⁰ Op.cit. 384 ff.

lyāk, pl. 3 *lyākar* to *lāk*- 'see' (present act. sg. 3 *lkāš* etc.); act. pl. 3 *cārkar*, mid. sg. 3 *cārkat* to *tār*k- 'let, let go' (present act. sg. 3 *tārñāš* etc.); act. pl. 3 *šārsar* to *kārs*- 'know' (present act. sg. 3 *kārsñāš* etc.); mid. sg. 3 *pārat*, pl. 3 *pārant* to *pār*- 'carry' (pres. mid. sg. 3 *pārtār* etc.); mid. sg. 3 *šālpāt* to *kālp*- 'find, get' (pres. mid. sg. 3 *kālpñātār* etc.); mid. pl. 3 *šākant* to *tsāk*- 'pull out' (present mid. sg. 3 *tsāknātār* etc.)

IIb shows the strong form of the root (*a*, *e*), to which the stem vowel *ā* is added before the endings (act. sg. 3 in *-s*), and likewise palatalizes the initial: act. sg. 2 *crankāšt*, 3 *crankās*, pl. 3 *crankār* to *trānk*- 'say' (present act. sg. 3 *trānkāš* etc.); act. pl. 3 *šepār* to *tsip*- 'dance' (present act. pl. 3 *tsipiñc*).

III is formed from the present stem, of which the final consonant is palatalized if possible, and adds the stem vowel *ā* (which is retained without ending in the act. sg. 3). This is the most common formation. Some seventy-five or more forms are listed by SSS; only a few examples will be cited: act. sg. 3 *ešā*, pl. 3 *ešār* to *e*- 'give' (pres. act. sg. 1 *esam*, 3 *eš*); act. sg. 3 *kātāñšā* to *kātk*- 'stand up, arise' (pres. act. sg. 3 *kātānkāš*), *keñā* to *ken*- 'call' (pres. act. sg. 3 *kenāš*); mid. sg. 3 *kropñāt*, pl. 3 *kropñānt* to *krop*- 'collect' (pres. mid. pple. *kropnmām*).

IV may be based upon the subjunctive stems in *-ās*- and in *-ñ*-: cf. mid. pl. 3 (*tā*)*kwāšānt* and a fragmentary *tākwāšā*..., apparently beside vbl. sb. *tākwāšlume*, which is then from the subjunctive stem of the causative verb to the root *tākw*- (inf. *tākwātsi* etc., cf. SSS 439 with no meaning given for the verb). As possible examples of imperfects from the subjunctive in *-ñ*- are cited (SSS 387) the following: mid. pl. 3 *tpukñānt*, act. sg. 3 *wātñā* and *tsākñā*. For only the last of these is either the meaning clear or the present formation attested, viz. pres. mid. sg. 3 *tsākāštār* 'glows'. The other attested forms are pret. mid. pl. *tsāksānt* and opt. mid. pl. 3 *tsāsintrā* (cf. SSS 481). The fact that the preterit middle is in *s* (Cl. III) makes it possible that the subjunctive is in *ñ*, e.g. act. sg. 3 **tsākñāš* etc.; but no such form is attested, and actually the optative *tsāsintrā* makes it more probable that the subjunctive is of a 'thematic' type, e.g. act. sg. 3 **tsākāš*, since for the *ñ*-subjunctive we would expect rather opt. **tsākñāintrā*. These imperfects then can very well have been formed (as already suggested by SSS 337) from unattested presents in *nā*, and belong rather to Cl. III. Indeed, there is no conclusive evidence for an imperfect based on the subjunctive stem in *ñ*, or in fact on one in *ās*, owing to the fragmentary nature of the quotable forms. We have then primarily three formations to examine: IIa, IIb, and III.

A fundamental assumption of this paper will be that like formations had in origin like grammatical meanings, and that differences of meaning between identical formations in the same language or in closely related languages are secondary. If we accept this assumption, I believe we can identify some at least of the Turfanian imperfects.

For type IIa, the key lies in the identification of Turf. impf. *lyāk*, *lyākar* (above) with Kuch. pret. *lyāka*, *lyākāre*, etc. This identity, which has been assumed over and over again, would seem indisputable;¹¹ Pedersen¹² adds to it

¹¹ Cf. SSS 385 fn. 1; Couvreur, *Hoofdzaken van de tochaarse klank- en vormleer* 66; Pedersen, *Tocharisch* 178.

¹² Loc.cit.

that of Turf. impf. pl. 3 *šārsar* and Kuch. caus. pret. sg. 3 *šārsa* (pl. 3 *šārsāre*) from *kārs-* 'know' (Turf. pres. sg. 3 *kārsnāš*, Kuch. *kārsanam*). He continues, however, to follow Wilhelm Schulze¹³ in the view that the Kuch. caus. *šārsa* results from contraction and is to be equated therefore also with Turf. (uncontracted) caus. pret. *šašārs*. As a result, he is faced with two problems: (1) the origin of the distinction in meaning, i.e. between non-causative imperfect and causative preterit; and (2) the reason for the presence of both contracted and uncontracted forms in Turfanian. As regards the first, Pedersen decides that the contrast between the contracted and uncontracted forms had in origin nothing to do with their grammatical (non-causative or causative) meaning, nor with dialectal differentiation; and as regards the second, he concludes that contraction took place properly only in longer forms. The resultant distribution in Turfanian and the elimination of uncontracted forms in Kuchean are due to analogy.¹⁴ With the view that the difference in grammatical meaning was not originally bound to the contrast in form, I heartily agree. But I have never been able to accept Schulze's view that Kuch. *šārsa* results from contraction and is to be equated to Turf. *šašārs*—a view which is all the less acceptable if (with Pedersen) we equate to it also the Turf. imperfect *šārsar*.

Schulze's theory of the identity of these forms was supported by the assumed parallelism with the development of the seventh class of strong preterits in Germanic: for example Goth. *haihait*, OE *heht* on the one hand, vs. OHG *hiaz*, OE *hēt* on the other. One cannot of course prove one improbable phonetic development by recourse to another just as improbable. Many students of Germanic, I among them, have long given up the contract origin of unreduplicated seventh-class preterits in North and West Germanic.¹⁵ As in the case of the Germanic forms so also in that of the Tocharian, I am convinced that the long-vowel preterits and imperfects in question have in origin nothing to do with the reduplicated forms. The coincidence in grammatical meaning between causative preterits like Kuch. *šārsa* : Turf. *šašārs*, or Kuch. *cāla* : Turf. *cacāl* (to *tāl-* 'lift') is purely accidental, as is so often the attachment of any particular meaning to a formal category. As to what these forms are from the Proto-Indo-European point of view, I can only reiterate my earlier opinion¹⁶ that the possibility of a connection with the 'long-vowel' perfects of Latin and Germanic should not be as categorically denied as it seems to be by Pedersen and others. The equation Turf. (imperf.) *pārat* (to *pār-* 'carry') with Goth. *bērum* etc., and even of Turf. *lyāk* (imperf.), Kuch. *lyāka* (pret.) with Lat. *lēgī*, is hard for me to reject. The comparison of these long-vowel perfects with the Hittite *ḫi*-conjugation makes it almost certain that these were originally Indo-European 'perfects' (and Indo-Hittite presents) denoting a present state or the result of an action performed in the past.¹⁷ I believe that the shift from the notion of

¹³ *Kleine Schriften* 239 ff. = *Sitzungsberichte der preuss. Akad. der Wissenschaften* 1924 166 ff.

¹⁴ *Op.cit.* 176.

¹⁵ Cf. Prokosch, *A comparative Germanic grammar* 176 ff., with literature.

¹⁶ Cf. *Lg.* 24.307-8.

¹⁷ On the identity of the Indo-European perfect with the Hittite *ḫi*-conjugation, cf. Sturtevant, *Comp. gramm. of the Hittite language*² 131 ff.

aspect to that of relative time, which took place in Tocharian as it did eventually in most of the other Indo-European languages as well, could quite easily transfer such formations to a developing category of imperfects (e.g. in Turfanian), or make them simple preterits to an originally iterative-durative or simply imperfective present system. This seems to be the original value of the *sk* formation in Indo-Hittite,¹⁸ and is still a common characteristic of some *sk* (§§) presents in Tocharian,¹⁹ though of course their predominant value in both dialects is causative. Kuchean preterits of Cl. II (caus.) like *šārsa* (= Turf. impf. *šārsar*, cf. above), *ñyārsa* (*nārs*- 'crowd'), *pyālka* (*pālk*- 'shine'), *myārsā-ne* (*mārs*- 'forget'), *šārkate* (*šārk*- 'excel'), *špyārta* (*spār*tt- 'turn'), *špyarkatai* (*spār*k- 'pass away'), *tsyālpāte* (*tsālp*- 'redeem'), probably have their long vowel by analogy with stems ending in a single consonant, like *cāla* and *lyāma*. It is most unlikely that the length of the vowel here reflects the original quantity. A parallel is shown in the West Germanic seventh-class strong verbs of the type of OHG *hialt*, *giang*, *fiang*, (after *hiaz*, *liaz*, etc.), where Gmc. *ē*² (< PIE *ēi*) cannot possibly be original. Kuchean Cl. II preterits in the diphthongs *ai* and *au* can reflect normal PIE perfects (without reduplication) with initial palatalization on the analogy of the 'long-vowel' type. Examples are *klyautka*, *klyautkate* (*klut*k- 'turn about', caus. 'make into'), *traika-ne*, *traikate* (*trik*- 'go astray'), *pyautka*, *pyautkare* (*pyut*k- 'come into being'), *raittate*, *raittānte* (*ritt*- 'be bound', caus. 'become bound'), etc. These Kuchean 'causative' preterits of Class II are then identical, so far as the vocalism of the root is concerned, with 'non-causative' preterits of the type of Kuch. *kāka*, *kakāte* = Turf. *kāk* (from pres. Kuch. *kwā*-, Turf. *ken*- 'call'), Kuch. *kamāte* = Turf. *kāmat* (*kām*- pret. stem to Kuch. and Turf. *pār*- 'carry'), Kuch. *kārpa* = Turf. *kārp* (Kuch. and Turf. *kārp*- 'descend'), Kuch. *paiykāte* = Turf. *pekat* (Kuch. and Turf. *pik*- 'write'), Kuch. *kraupāte* = Turf. *kropat* (Kuch. *kraup*-, Turf. *krop*- 'gather'), etc.²⁰ The only difference is the lack of initial palatalization in both dialects. No palatalization is to be expected, of course, if these forms are reflexes of the PIE normal perfect singular of the *e/o* series (Gk. *λέλοιπα*, *oīda*, etc.). It would appear then that the forces of analogy worked in two different directions: on the one hand, the causative preterits and non-causative imperfects all assumed the palatalization after the model of the 'long-vowel' type (*lyāka*, *lyāk*, etc.); on the other, the non-causative preterits rejected it where one might expect it to have occurred originally, as in *kāka*, *kāk*, *tāka*, *tāk*. It hardly seems possible to assume that these latter are all reflexes of 'long-vowel' perfects of other series (like Goth. *stōþ*, *stōþum*, Lat. *fōdī*, *scābī*). The principle of 'vowel balance' or vowel weakening in connection with the primitive accent position and number of syllables²¹ seems to work more uniformly in the Kuchean pret. Cl. Ib than in Cl. II; but this can hardly be significant, especially in view of the great irregularity of its application in all categories of forms. A few examples will make this principle clear: Kuch. Cl. Ib act. sg. 3 *kāka* but mid. *kakāte*, mid. sg. 1 *kāmmai*, but sg. 3 *kamāte*, act. sg. 3 *kārpa* but sg. 1 *karpāwa*, etc., as opposed

¹⁸ Cf. Sturtevant, op.cit. 129-30; Couvreur, *Revue des études indo-européennes* 1.89 ff.

¹⁹ Cf. Krause 83.

²⁰ These are Krause's Class Ib, but SSS Ia.

²¹ Krause 10 ff.

to Cl. II act. sg. 3 *cāla* and sg. 1 *cālawā*, act. sg. 3 *kyāna*, sg. 1 *kyānawā*, act. sg. 3 *lyāma*, mid. 3 *lyāmate*, and the like. The alternation of *ā* in dissyllabic forms with *a* in trisyllabic ones is thus the rule in Kuchean Class Ib but not in Class II. That the stem vowel (between root and ending) was originally long in the latter class also is clearly shown by forms like *myārsā-ne* with suffixed pronoun (*mārs-* 'forget'), *raittānte* beside *raittante* (*ritt-* 'be joined'), *tsyālpāte* (*tsālp-* 'be redeemed').

Let us turn now to Turfanian imperfect IIb. This, as we have already seen, is represented by only four forms, all active: sg. 2 *crāṅkāšt*, 3 *crāṅkāš*, pl. 3 *crāṅkār* from *trāṅk-* 'speak', and pl. 3 *šepār* from *tsip-* 'dance'. These are identical in formation with preterits of Class III, e.g. act. sg. 3 *ñakās* (*nāk-* 'disappear'), *lyepās* (*lip-* 'remain, be left'), *lyokās* (*luk-* 'light up'), pl. 3 *caṅkār* (*tāṅk-* 'hinder, obstruct'), *crakār* (*tār-* 'let, release'), *šarkr-ām* (*kār-* 'bind'). It follows, if my fundamental assumption is correct, that they are in origin the same formation. This is also the Class III 's-preterit' of Kuchean, cf. sg. 3 *neksa* (: Turf. *ñakās*), *lyauksa* (: Turf. *lyokās*), etc. The coincidence of forms in the two dialects is extremely close.²² Both dialects show both causative and non-causative values, and both show both palatalized and non-palatalized initials. In Turfanian the palatalization is largely confined to the active, and therefore agrees with that of the imperfect IIb (SSS) which I identify with it, although only active forms of the latter are attested.

As regards the history of the s-preterit (and hence of the Turfanian s-imperfect IIb), I am inclined to agree with Krause²³ that it is probably of mixed origin. The Indo-European s-aorist furnished the endings with -s, e.g. act. sg. 3 Kuch. *neksa*, *lyauksa*, *lyautsa*, Turf. *ñakās*, *lyokās*, and the entire Kuchean middle (*neksate*, *lauksāte*, *lyutsāmai*, *lyutstsatai*, *lyutstsante*), but in Turfanian only the middle of the second type (e.g. *rise*, *risāte*, *risāt risānt*, cf. SSS 376).²⁴ The Indo-European perfect, on the other hand, probably lies back of the radical in Kuch. act. 1 *nek-wa*, 2 *nek-asta*, 3 *nek-sa*, which must represent PIE o-grade forms, (cf. Lat. *nocuī*, *nocuisti*, *nocuīt*). Since the *e* < IE *o* in these forms is probably to be equated with the Turf. *a* in *ñakās* (cf. Kuch. *keme* = Turf. *kam* 'tooth' : Gk. *γόμεφος*, OIcel. *kambr*, etc.),²⁵ the Turfanian palatalization in such forms is clearly not original. It could arise, however, in forms of the s-aorist, which regularly had the e-grade (cf. Gk. *ἔδειξα*, *ἔπεισα*, etc.). Hence such forms as Turf. preterit *lyepās*, *lyokās* (with PIE *ei*, *eu*) and imperfect *šepār* are probably correct, and so are the Kuchean s-preterit forms with initial *ly* cited above.²⁶ It is very doubtful that the initial palatalization in the Turf. imperfect *crāṅkāšt*, *crāṅkāš*, *crāṅkār* (to *trāṅk-* 'speak') is original; it is of course parallel to that observed in the pret. pple. *caccriku* (beside *tatriku*) to *trik-* 'be confused, err', which may likewise be

²² Cf. Krause's list 181 ff.

²³ Op.cit. 180.

²⁴ The question of the identity of the 2nd sg. ending Kuch. -*sta*, Turf. (*ā*)*št* with Lat. *isti* need not concern us here; cf. Krause, loc.cit.

²⁵ Lg. 14.28-9; Pedersen, *Tocharisch* 219.

²⁶ Krause's remark (180) that palatalization is unknown in the Kuchean preterit of Class III is in general true, but initial *ly* occurs.

'functional' rather than original. The palatalization of consonants across an intervening 'unpalatalizable' consonant is of course not unknown in Turfanian; thus—across *w*: Turf. *špām*, Kuch. *špane* 'sleep' (: Skt. *śvapnas*, OE *swefn*); Turf. *šar*, Kuch. *šer* 'sister' (: Skt. *śvasar-*, Goth. *swistar*);—across *p*: Turf. *špin-ac* (dat.) 'peg' (: Lat. *spīna*); Turf. *šašpānku* caus. pret. pple. from *tšpānk-* (SSS 484).²⁷ But there is no good reason to assume that this was not also analogical palatalization and that the form is therefore an original perfect.

It remains to discuss the most widely represented class of imperfects in Turfanian, Class III—those derived from the present stem by addition of the stem vowel *ā* (< PIE *ē*) with preceding palatalization. That some of these are identical with aorists has been already noted by Pedersen:²⁸ 'Einige *ē*-Imperfekte haben die Geltung als Präterita erworben. Das hängt mit der Wortbedeutung zusammen.' He cites the three Turfanian roots *klyos-* 'hear', *wles-* 'perform, practice', and *pās-* 'keep, observe'. According to SSS 386, the forms act. sg. 1 *klyoṣā*, pl. 1 *klyoṣāmās*, 3 *klyoṣār*, mid. sg. 1 *wle[še]*, 2 (*pā*)*ṣāte*, 3 *wleṣāt*, *pāṣāt* are imperfects, but the same forms are cited also (381–2) as preterits of Class IVc. To these must be added (*wi*)*nāṣār* to *winās-* 'honor' (likewise cited in both places by SSS), and possibly *wināṣā-ṇ* (cited as imperf. SSS 472). Again the identity of the forms—here within one dialect—is the key to their identification in the total system of imperfects and preterits in both dialects. In Kuchean the corresponding forms are of course preterits only: sg. 1 *klyauṣāwa*, 3 *klyauṣa*, pl. 3 *klyauṣāre*, etc., mid. sg. 2 *paṣṣatai*, 3 *laṃṣṣāte*, pl. 3 *laṃṣṣānte*, *lāṃṣānte*, all classed by Krause as Class Ib (= SSS Ia); and act. sg. 3 *w.nāṣṣa-me*, pl. 3 *wināṣṣā(re)*, *wināṣṣar-ne*, classed by Krause as IV, since the latter is clearly derived from the present stem in *-sk-*. Actually the formation of the pret. *w.nāṣṣa*, beside the present *wināsk-*, is identical with that of *klausa*, *paṣṣatai*, *laṃṣṣāte*, and all the other preterits which Krause classed as Ib and which show as their characteristic the palatalization of the final radical consonant or consonant group, e.g. *aklyamai*, *aklyyatai* from *ākl-* 'learn'; *akṣāwa-me*, *akṣāsta* from *āks-* 'announce' (pres. *aksaskau*); *kaccāre* from *kātk-* 'rejoice' (pres. *kātkau*, *kāccām*).²⁹ This formation has no parallel among the corresponding Turfanian preterits (SSS cl. Ia), except in the preterit-imperfect forms (SSS IVc) cited above. On the other hand it has been retained and enlarged as the regular imperfect when formed from the present stem: *eṣā* to *e-* 'give' (pres. *esam*, etc.), *karyā* to *kar-* 'laugh' (pres. *karyeñc*), *kātāñṣā* to *kātk-* 'arise' (pres. *kātāñkāṣ*), etc.; cf. the survey given above.

As to the ultimate Indo-European provenience of this formation, Pedersen³⁰ has already connected it with the *ē*-verbs of the other Indo-European languages, which, as is generally conceded, expressed states or conditions rather than actions; cf. Lat. *pendeō*, *jaceō*, *taceō* (= OHG *dagēn*). In general, the verbal stem in *ē* gives non-present forms in other IE languages (specifically in Greek, Baltic, and Slavic) which are frequently parallel to present formations in *yo/ye*; cf. the Gk.

²⁷ Lg. 21.22, 23.

²⁸ Op.cit. 180.

²⁹ Krause 167–8.

³⁰ Op.cit. 179–80.

second aorists ἐχάρην, ἐμάρην, ἐφάρην, beside the presents χαίρω 'rejoice', μάλωμαι 'rage', φάλω 'show'; Lith. preterits *smirdėjau*, *sėdėjau*, beside presents *smirdžu*, *sėdžu* 'sit'; OCS aorists *smirděchŭ*, *mĭněchŭ*, *viděchŭ* beside presents *smĭrĭdŭ* 'stink', *mĭnjŭ* 'think', *viždŭ* 'see'.³¹ In Latin and Germanic, however, the stem in *yo* may form an originally transitive verb, parallel to the originally intransitive present in *ē*; cf. Lat. *jaciō* 'throw' beside *jaceō* 'lie (thrown)'. In Germanic the original difference is usually obscured: Goth. *hafjan*, OHG *heffen* 'lift' = Lat. *capiō* beside Goth. *haban* (*habaiþ*), OHG *habēn* (*habēt*) 'have' (i.e. originally the state of lifting); Goth. *hatjan* 'hate': *hatan*, OHG *hazzēn* 'id.'

It seems clear to me that these imperfects (and/or preterits) of Turfanian and the preterits of like formation in Kuchean are to be equated (disregarding differences of ending) with the second aorists of the type of Gk. ἐχάρην. This view is considerably strengthened by the fact that in Tocharian also we find them co-existing with present stems in *yo/ye*. A conspicuous example is Turf. *kar-* 'laugh', act. pres. pl. 3 *karyeñc*, impf. sg. 3 *karyā*. In Kuchean, unfortunately, only present forms (act. pl. 3 *keriyem*, mid. pple. *keriyemane*) and the pret. pple. *kek(e)ru* are attested. I believe, however, that this verb can be confidently equated, sound for sound and form for form, with Gk. χαίρω, ἐχάρην.³² Other examples of *yo/ye*-presents beside *ē*-preterits are seen in Turf. *kāl-* 'stand, be', pres. mid. sg. 1 *kālymār*, 3 *kālytār*, pl. *klyantrā*, imperf. mid. sg. 3 *klyāt*, pl. 3 *klyānt*; and *malyw-* 'crush', pres. act. sg. 2 *malywāt*, impf. act. sg. 3 *malywā*. Both verbs are also *yo*-presents in Kuchean (pres. mid. pl. 3 *klyentār*, etc., sg. 3 *melyan-me*, pl. 3 *melyem*). Again no preterit is attested in Kuchean to either of these stems (the preterit of *kāly-* is from the suppletive *stām-*), but the preterit active to the Kuchean *yo*-present *pānn-* 'stretch', pres. mid. sg. 3 *peññatrā*, is clearly of the expected formation: sg. 3 *piñña* (mid. sg. 3 *pānnāte*, however, without palatalization).³³ To the corresponding Turf. *panw-* (also *yo*-present, act. sg. 3 *pañwās*), no imperfect of any sort is attested. Another possible *yo*-present is Turf. *me-*, *mew-* 'tremble', pres. act. sg. 3 *meş*, pl. 3 *meyeñc*, with *ē*-imperfect *meyā*. The Kuchean correspondent is *miw-*, pres. sg. 3 *miwām*, pret. act. sg. 3 *maiwa*, mid. sg. 3 *maiwāte*. But in the absence of a clearly palatalizable consonant and a convincing etymology, there is no real evidence.

Two other Kuchean preterits seem to me to belong definitely to this formation: *karşsa* (*karşş-* 'shoot') and *kālpawā* (*kālyp-* 'steal'), inasmuch as they may show the expected reduced-grade vocalism (cf. ἐχάρην above). Here, however, no present forms are attested. Kuch. pret. *camyāwa*, *campya* (to present *campām* 'can') and pret. pl. 2 *memyas* (present not attested) probably show analogical 'secondary palatalization'.³⁴

³¹ Cf. Brugmann, *Grdr.* 2 2.3.158 ff.; Buck, *Comp. gramm. of Gk. and Lat.* 269.

³² In such a case then PIE *tr* (or *tr* or *r* before vowel, however we write it) fell together in its development with PIE *o* and *a*. The etymology that I proposed in *Lg.* 14.29 (connection with Lat. *garrĭō* 'chatter'), is therefore to be abandoned, as is also that of Van Windekens, *Lex. etym.* 37.

³³ Cf. Krause 166 Anm. 1.

³⁴ Krause 165 (§164) compares these last-mentioned preterits, *karşsa*, *kālpawā*, *comyāwa*, *memyas*, with Baltic preterits like Lith. *tempiaũ* (to *tempĭũ*, *tempĭti* 'stretch'), *tapiaũ* (to *tampũ*, *tapti* 'become'), as formations in *-yā*. I prefer the standard view (as in Brugmann,

This formation *ā* has of course been extended in Turfanian as an imperfect to all sorts of presents, including those in *-nā* (*klišnā*, pres. *klišnāṣ* 'sleeps'), those in *-nās* (*kumṣā*, pres. *kumnāṣ* 'comes'), and particularly those in *-s-*, primary and causative (e.g. *eṣā* to *e-* 'gives', *kātkṣāt* to *kātkṣātār* caus. to *kātk-* 'rejoice'), as well as to denominatives in *iññā-* (e.g. *tuñkiññā*, inf. *tuñkiññsi*, ppl. (*t*)*uñkiññant-*). As a preterit formation we find the identical suffix also in Class IVb (SSS 380 f.): *ākṣiññā*, *okṣiññā* (from *āks-* 'teach' and *oks-* 'grow', presents *ākṣiṣ* and *okṣiṣ*). The same formation is found (only as a preterit, of course) in Kuchean, e.g. *kwipeññate* (to *kwip-* 'be ashamed', pres. *kwipeññentār*). Pedersen has already identified the formation as a *yo*-suffix added to an *n*-stem, and has compared Skt. *iṣanyati* 'incites'.³⁵ Another parallel is OHG *giwahanen* 'mention', PIE **wokⁿ-yō*. The latter etymology is important here, since it shows the present formation to be expected beside the Tocharian preterit of the verb 'say': Turf. act. sg. 1 *weñā*, 2 *weñāšt*, 3 *weñā* = Kuch. *weñāwa*, *weñāsta*, *weñā*. The present of this verb, however, is furnished in Kuchean by an *sk*-formation (act. sg. 1 *weskau*, 3 *wesṣām*) and in Turfanian by a suppletive verb (act. sg. 1 *trānkām*).

The remaining Turfanian imperfects, those presumed to be derived from subjunctive stems (Class IV, SSS 386), are too inadequately attested for detailed discussion. Purely on the basis of form, however, they too are identical with known preterits: (*tā*)*kwāṣānt* (mid. pl. 3) and fragmentary *tākwāṣā..* are to be identified with Turf. pret.-impf. *wināṣā-m*, (*wi*)*nāṣār*, Kuch. pret. *wināṣṣa*; Turf. *tpukñānt*, *wātñā*, and *tsākñā* are identical with *weñā*, etc., above.³⁶

In conclusion, I believe that the evidence brought forth here supports the view expressed at the outset: that the Kuchean imperfect of optative origin is to be attributed to Proto-Tocharian, and that in Turfanian it was replaced by forms originally identical with those which came to be used as preterits. This identity is for the most part demonstrable within Turfanian itself, but is even more obvious in the light of comparisons with Kuchean. These 'preterits' of Proto-Tocharian are probably derived from Proto-Indo-European perfects and aorists. The eventual developments of some of them into Turfanian imperfects must be explained by the meaning of the verbs themselves; but their origins must be identified exclusively on the basis of form.

Grdr. 2.3.176-7), that these Baltic forms are also *ē*-stems, not *yā*-stems; cf. sg. 2 *tempeī*, 3 *teñpē*, pl. 1 *teñpēme*, 2 *teñpēle*. They are then indeed comparable to our Tocharian formation, but only in the way that I have indicated.

³⁵ *Tocharisch* 170.

³⁶ Couvreur (*Hoofdzaken* 66) may be right in suspecting that we are actually dealing here, at least in part, with preterits and not with imperfects. In fact, while throughout this paper I have accepted the classification as imperfect or as preterit as I found it in SSS or Krause, I am not at all sure that when all texts are eventually analyzed such a classification will hold for Turfanian.

PA, ΔΕΔΑΕ, ΔΑΣΤΣ, AND THE SEMIVOWELS

HENRY M. HOENIGSWALD

University of Pennsylvania

Given the following groups of corresponding phonemes or phoneme sequences in Sanskrit, Greek, and Germanic: (1) $v/\text{f}/w, r/\text{p}/r, n/\text{v}/n$, (2) $u/v/u, \tau/\text{pa} \sim \text{ap}/ur, a/\text{a}/un$, and (3) $w/v(\text{f})/(uw), ir(ur)/\text{ap}/ur, an/\text{av}/un$: the segments of sound lying behind group 1 were once in complementary distribution both with those behind group 2 and with those behind group 3. For the older stages of Indo-European, therefore, one needs to reconstruct only three entities: w, r, n (and likewise y, l, m for other sets not listed here), each with three positional variants of varying syllabicity. The argument has been most fully presented by Franklin Edgerton.¹ In connection with that argument there has been some discussion whether the syllabic element in group 3 above should or should not be written v .² Assuming for the moment that v has to be reconstructed in the neighborhood of stops to account for *quattuor*, *πiουρες*, and the like, the dilemma resembles that posed by the $[k]$ of *length* in the description of some types of English: if we are impressed by its similarity to other $[k]$ -like segments in environments where it is not automatically present, we say that there are some environments in which $[k]$ is automatic but is phonemically $/k/$ nevertheless; but if we are more impressed by the automatic nature of its presence in *length*, we seize upon any phonetic detail that may distinguish this $[k]$ from other $[k]$ -like sounds elsewhere and see in the sequence $[\eta k]$ merely the allophone of $/\eta/$ before $/\theta/$.³ If v is not recognized in *quattuor*, the problem hardly arises in the case of the nasals and liquids. For $(i)y$ and $(u)w$ it still exists, in a somewhat different form; but this question is here left out of account.

Even granting the possible need for v in the phonemic inventory of Proto-Indo-European, Edgerton refuses to recognize it before liquids and nasals under the conditions that gave rise to group 3 above—that is, after two nonsyllabics (or vowel length or pause plus nonsyllabic) and before a syllabic. Of the two alternative solutions mentioned for the $[k]$ of *length*, he prefers the second. The so-called converse of Sievers' Law, or the treatment of two like semivowels coming together at a morphological seam, may not absolutely exclude the use of v (since one could still say that the first of the pair is lost after a light syllable and replaced by v after a heavy syllable); but the interpretation of the data is incomparably more elegant without the use of v . In fact, semivowels turn out to behave simply like consonants. As **es-si* 'thou art' became **esi* within Indo-

¹ *Lg.* 10.235-65, 19.83-124. See also G. S. Lane, *JEGP* 50.522-8.

² See E. H. Sturtevant, *Lg.* 19.293-312.

³ Or of $/\theta/$ after $/\eta/$. It would still have to be shown in what sense the $[k]$ of *strength* is more automatic than the weak-stressed vowel of *(he) longeth* or the like: both sounds occur between $/\eta/$ and $/\theta/$ (and so do countless other sequences, especially longer ones); zero does not.

European (Skt. *asi*, Gk. *ei*), so Skt. **cakr-re* appears as *cakre*. Whether we have *r* or *ir* in Sanskrit depends automatically on the weight of the preceding syllable (light in this example), not on the number of *r*'s in some earlier stage of the form.^{2b}

The sandhi variants that arise 'when a semivowel before a vowel occurs either at the beginning of a word or preceded only by a single consonant ... depending on the nature of a preceding word final' are still close to their original distribution in the Rigveda, although generalization has begun.³ Other languages are said to show traces of the original distribution in their suffixes (Gk. *κοινός*, *κάμνω*, but *ἄγριος*, *λανθάνω*) but not in sentence sandhi, where only generalized doublets survive (*δύω* but *δώδεκα*, without regard to the context). On the borderline between morphology and syntax, however, is the case of the Greek post-positive particle *ἄρα*, *ἄρ*, with its enclitic by-form *ῥα*. This case merits some attention.

Ἀρα may be an adverb like *τάχα* or *μάλα*; in that case we would be dealing with a semivowel (*r*) before a vowel. *Ἀρ*, rare except in *αὐτάρ*, *γάρ* (not counting the elided *ἄρ'* before a vocalic initial), may well show apocope on the model of *πάρ* and *ἄν*. Leaving *ἄρ* aside, we find that *ἄρα* and *ἄρ'* seem to occur quite freely in Homer. *Ῥα*, on the other hand, does not. Of 91 instances of the unelided form in the first twelve books of the Iliad, only 8 are found after a heavy word-final syllable (e.g. A 430 *τὴν ῥα βίη ἄεκοντος ἀπηύρων*), and of 95 instances of the elided *ῥ'*, not one occurs in this position; in fact, there is only one such occurrence in the two epics together (N 85 *τῶν ῥ' ἅμα ἀργαλέω καμάτῳ*). Of course, elided *ῥ'* after a heavy syllable is metrically without effect, while *ῥα* makes a difference. That the avoidance of *ῥα* and *ῥ'* after heavy syllables is not an accident is confirmed by a countercheck: of 32 instances of *ἄρα* and *ἄρ'* in one book of the Iliad (M), as many as 9 follow an unelided heavy syllable such as *ἐξ*, *ὦς*, or *ἄς*.⁴

Grammont has dealt with *ῥα*, though in a somewhat different sense to be discussed below.⁵ Curiously enough, in explaining the variant forms as going back to possible sandhi doublets, he seems to be unaware that he need not speak of this theory as mere speculation, and that the statistical data strongly favor it. He was perhaps misled by an old article by Hiller, whose whole aim had been to stress the fact that *ῥα* is somewhat restricted to occurrence after monosyllables, and to explain away, more or less plausibly, examples to the contrary.⁶ Hiller did not see the significance of the scarcity of heavy syllables before *ῥα*.

This may cast a different light on some of Grammont's other arguments, which have so far evoked only very restrained enthusiasm.⁷ For Grammont, *ῥα* represents **r*, not **ra*. *Ἀρ* also comes from **r*, while *ἄρα* is a blend of the two

^{2b} Somewhat differently Sturtevant, op.cit. 312.

³ *Lg.* 19.91.

⁴ *Ῥα* and *ῥ'* occur freely enough after long diphthongs.

⁵ *Phonétique du grec ancien* 285-6.

⁶ *Hermes* 21.563-9.

⁷ See E. Risch, *Museum helveticum* 6.244.

forms. Here he is in agreement with Brugmann, who adduced Lith. *ir*, with strong semantic support.⁸ Such a view is of course entirely possible. If it is correct, the distribution of *pa* in Homer (as described above) gives substance to an old idea⁹ taken up by Grammont, namely that the difficult vacillation between *ap* and *pa* for **r* (i.e. for that allophone of the semivowel which is found, roughly, between pauses or consonants) is also due to the structure of the preceding syllable—in other words, that something like Sievers' Law holds good in Greek not only if a vowel, but also if a consonant (or a pause) follows. As a matter of fact, most neuters in *-ap* from **r* have, or had prehistorically, a heavy syllable before the **r*: *oĩθap*, *ēap* < **wēsr(t)*. On the other hand, *ὑπόδρα* from **ὑποδρακ(r)* had a light syllable there. Doublets like *καρδία* : *κραδία* would be due to sandhi, with complete generalization of both variants. Initial **r* seems to be always represented by *ap*, a fact which may be ultimately related in some obscure way to the presence of a laryngeal before it.¹⁰

Those who have thought of explaining the difference between *ap* and *pa* as a result of syllabic rhythm seem to have envisaged it as a specifically Greek development: when **r* was given up as a homogeneous syllabic segment, it is supposed to have gone either to *ap* or to *pa*, depending on its environment. But the development may just as well have been in the opposite direction. If so, an automatic vowel segment developed before the semivowel where two consonants preceded, but after the semivowel where only one consonant preceded: *akt₁rt* but *atr₁t*, to use the customary notation.¹¹ In Greek the distinction continues, with the automatic vowel becoming phonemic through merger with other vowels in the language and also through the generalization of sandhi variants. In Sanskrit, *r* and *r₁* fall together in *r*. With regard to the liquids, Sanskrit behaves as all the IE languages do in the case of *y* and *w*: nowhere do we find any distinction made between the semivowels in /aktwt/ and /atwt/; both positions show the vowel *u* (*aktut*, *atut*).

What about the nasals? In Indo-Iranian and in Greek, the allophones *ŋ* and *ŋ₁* (i.e. the phonemes /n/ and /m/ between pauses or consonants) appear as *a/a₁*; in the other languages they fall together with sequences of vowel plus *n* or *m*. Yet it is not impossible that here too Greek may have preserved a trace of the allophonic difference between *n* after a heavy syllable and *n₁* after a light one. Here, in fact, may lie the solution of the controversy which has attached itself to the etymology of *δαρός* 'thick'. There is no good reason to doubt that this word has the *s* of Lat. *dēnsus*, except that one would have to assume that the so-called *ŋ* of **dŋsus* was not yet *a* at the time that *s* was lost between vowels.¹² And yet

⁸ *Verhandlungen d. Sächs. Ges. d. Wiss., Phil.-hist. Kl.* 35.36-70.

⁹ See Brugmann himself, *op.cit.* 67.

¹⁰ The relation between semivowels and laryngeals needs more investigation. A laryngeal would have to be expected in *apa* as well, especially if it belongs with *ἀπαπλωκω* 'join' etc.; see Boisacq s.v. And see W. P. Lehmann, *Lg.* 27.13-7.

¹¹ F. Edgerton, *Lg.* 19.83-124 passim. We use italics for a notation in which the allophones of the semivowels are distinguished, slant lines where they are not.

¹² W. Schulze, *Kleine Schriften* 116-7. *Πράσον* 'leek' proves little for *r* preceding *s*. Ernout-Meillet³ s.v. *porrum* consider it a Mediterranean loanword in both Greek and Latin.

δέδαε 'he taught' is almost certainly from **ded_hse(t)*. The material is naturally scanty. Schwyzler lists on the one hand δασύς and ἄσις 'mud' (cf. Skt. *ásita*), and on the other δανλός 'dense', which one would like to link with δασύς; and finally ἦα 'I was' from **ēs_hn*.¹³ Of these, only δέδαε and ἦα are safe from sandhi variation.¹⁴ If there was an automatic vowel, it cropped up after the *n* in **ded_n.set* and before the *m* in **ēs_n.m*; in both cases the *s* came to stand between vowels and was lost. On the other hand, if an initial nasal developed its automatic vowel in front of its consonantal component rather than after it (as the development of *r* suggests), we expect **n_si-* (whence ἄσις) and not **n_s.si-*. Then δανλός would represent the sandhi doublet generalized from positions after a world-final short vowel (**...a dn_s.ulos*), with its counterpart δασύς similarly generalized from other positions (**...at d_s.nsus*). All this is offered with the utmost caution. The few ascertainable facts at least permit the interpretation that has been sketched here.

We have in fact been debating the merits of a different notation for the syllabic allophones of the IE liquids and nasals. In one sense the difference is purely notational and nothing more, in that Edgerton's *o* for group 2 (and *r l n η* for group 3) has merely been replaced by *o* and that this *o* is considered just as automatic as the mark of syllabicity. In another, a phonetic sense, the difference has more reality. We seem to see a linguistic structure in which a nondistinctive vocalic element turns up before consonants and pauses under particular conditions: before a liquid or nasal if two consonants (etc.) precede, and after a liquid or nasal if only a single consonant intervenes after the last preceding vowel. As the Greek evidence seems to show, this is true both for group 2 (*o*, *r_o* = '*r*'; *n_o*, *n_o* = '*η*') and for group 3 (*o* = '*rr*', *n_o* = '*ηη*')—in other words regardless of whether a consonant or a vowel follows in turn. (In the latter case the automatic segment *o* can occur only after a heavy syllable or before the nasal or liquid.) Under the conditions that call for sounds of group 3, the daughter languages have kept the vowel segment in place, merely letting it fall together with one of the already existing vowel phonemes. Before consonant or pause, where sounds of group 2 are called for, only Greek may preserve traces of the location of the vowel segment, more clearly in οἶθαρ and ὑπόδρα than in ἄσις and δέδαε; in Indo-Iranian, *o* and *r_o* gave *r*, and *n_o* and *n_o* both fell together with *a* (as in historical Greek). The other languages also allowed the two variants to coincide with existing or possible sound sequences, with special problems arising here and there.

The semivowels proper, /y/ and /w/, may well have had a similar history, but it must be more remote. No trace of a distinction between *y* and *y_o* seems to exist anywhere: *i* (*iy*) and *u* (*uw*) appear uniformly in all the languages.¹⁵

Edgerton quotes with sympathetic interest a view according to which *o* itself, in the environment of a stop or spirant (or pause), may at one time have been

¹³ *Griechische Grammatik* 1.307.

¹⁴ It is true that the reduplicated forms all come from the Odyssey.

¹⁵ Even advocates of *o* believe that *o* before *y* and *w* became respectively *i* and *u*; see Sturtevant, *op.cit.*

automatic in its occurrence.¹⁶ Scholars who incline to this view have had to admit their inability to state the precise conditions under which such consonant sequences develop the segment ʔ. If this could be shown to have taken place after every cluster of two consonants (provided no vowel followed), all would be simple: ʔ and ʔ would be identical, and neither would be phonemic. There are some indications pointing in that direction, especially if the laryngeals are taken into account. But we cannot hope to reconstruct this uncomplicated state of affairs in all rigor, even for the oldest attainable stage of Indo-European.

¹⁶ *Lg.* 10.264. Perhaps it should be pointed out that our ʔ, being automatic, has nothing to do with similar notations used for a true reduced grade, where the location of the remnant vowel is determined by the location of the full-grade vowel in other forms of the paradigm. It is Edgerton's contention (and of course we agree) that such a true reduced vowel cannot be found in the neighborhood of semivowels. (See also J. Kurylowicz, *Études indo-européennes* 1.80-1 and 255-7.)

THREE PUZZLES IN THE LANGUAGE OF THE ILIAD

GEORGE MELVILLE BOLLING

Ohio State University

It is an honor to be invited to contribute to a number of *LANGUAGE* dedicated to Franklin Edgerton, and it is a pleasure for me to accept the invitation. For that purpose I have selected three notes dealing with the language of the *Iliad*, which are held together by the fact that all three end with problems of the sort that Edgerton has dealt with so successfully in his reconstruction of the *Panchatantra* and in his edition of the *Sabhāparvan* of the *Mahābhārata*—problems of textual criticism.

1. For the first of these I shall revert to a note published in the earliest days of our friendship: The etymology of *ΟΔΥΣΣΕΥΣ*, *AJP* 27.65-7 (1906). It started from Kretschmer's idea that, because of the dialect forms *Ὀλυντεὺς*, *Ὀλυσσεὺς*, and Latin *Ulixes*, the priority of the λ-forms must be recognized. *Ὀδυσσεὺς* is a 'popular etymology' for *Ὀλυσσεὺς*, made under the influence of *ὀδύσσομαι* and thus seeming to get sense out of what seemed senseless.

The attempts to go in the reverse direction, by assuming a change $\delta > \lambda$, have not succeeded. Brugmann-Thumb⁴ 126 said: 'die Belege für λ aus δ (Solmsen, KZ 42.211 f.) sind unsicher'; similarly Schwyzler 209: 'Ὀλυσσεῖδαι lat. *Ulixes* beweisen nicht spontanen (auch nicht dissimilatorischen) Übergang von δ in λ (so Solmsen, KZ 42.211 f.)'. Buck, *Comp. gram.* 123, refuses to class *Ulixēs* with *lacrima*, *lingua*, *oleō*, *lēvir*: 'for the l is attested in Greek variant forms of *Ὀδυσσεὺς* as *Ὀλυσσεὺς*, *Ὀλιξής*', thus implying agreement with Brugmann-Thumb and Schwyzler.

Phonology can take us one step further. Because of the variation $\sigma\sigma : \tau\tau$, the sound comes from either *κγ* or *χγ*. I choose the former and set up an earlier **Ολυκγος*. It is evidently a short name, and the problem is to find the full name from which it was derived. I suggested **Αυτολυκγος* a derivative from *Αὐτόλυκος*, the name of Odysseus' maternal grandfather. It was he who, according to τ 395-466, picked the name for his newborn grandson. It is there given, of course, in the epic form with δ; and Autolycus is made to explain his selection of it with the usual popular etymology. The explanation is not very satisfactory, and one is left with the impression that the story is a recasting of an earlier version in which Autolycus named the child after himself **Αυτολυσσος*.

My notion that *Ὀλυσσεὺς* is the short form of this name was advanced with the full consciousness that it was open to serious objection. Solmsen, KZ 42.208 (1908), declared it to be assuming an 'innerhalb des griechischen Namensystems beispiellose Art der Kürzung, bei der vom ersten Gliede gerade nur der Stammvokal übrig geblieben wäre.' I appreciated the force of his objection, and for years the problem seemed to me to be on a dead center: Solmsen's efforts to establish a change $\delta > \lambda$ needed to be abandoned;¹ while no progress had been

¹ Cf. the opinions quoted above, and add Kretschmer, *Glotta* 3.331-3 (1912).

made from starting with 'Ολυσσεύς. Recently I have come to see that 'beispiellos' was not the right word for Solmsen to use.

In the Catalog of the Nereids, Σ [39-49], one of the goddesses is named Καλλιάνειρα, another 'Ιάνειρα, a third Καλλιάνασσα, and another 'Ιάνασσα. These are good examples of the kind of shortening that I assumed.

It might be objected that the Catalog, as a plus (interpolated) passage, is not competent to testify. As a matter of recensio the Catalog is on a doubtful footing. There is a great disturbance of the tradition: omission in the Argolike, athetesis by Aristarchus and his predecessor(s). Perhaps, however, this is only a haplographic blunder, for the trap is great: Σ 38 πᾶσαι ὅσαι κατὰ βένθος ἄλδος Νηρηίδες ἦσαν : [49] ἄλλαι θ' αἱ κατὰ βένθος ἄλδος Νηρηίδες ἦσαν. Philologists have not agreed whether this Catalog is earlier or later than that in Hesiod. Surely it cannot be placed later than the 5th century; and for the present purpose that is early enough.

2. There seems to be competition between ἐπὶ + gen. and ἐπὶ + dat. in situations such as might lead in English to phrases like *to lie on the ground*, *to lay something on the ground*. Examples are: οἱ δ' ἐπὶ γαίῃ | κείμενοι Δ 161, τὸν δ' ἐπὶ γαίῃ κείμενον P 85, ἐπὶ γαίῃ | κείτο ταθείς N 654, Φ 118; τὰ μὲν κατέθεντ' ἐπὶ γαίῃ Γ 114—and with substantially equivalent verbs ἐπὶ γαίῃ | κάππεσε Π 310, 413 τὸ μὲν ἄψ ἐπὶ γαίῃ | πρόσθεν Μηριόναο πάγῃ ποδός Ψ 876; ἐξεράνησσ' ἐπὶ γαίῃ P 58. There is one example of the genitive τὸ δ' ἡμῖν κείτ' ἐπὶ γαίης N 565. The difference between ΓΑΙΗΙ and ΓΑΙΗC is slight. Variants are reported at Π 310, P 58, Φ 118, and also at N 565. It is difficult to see why γαίης was preferred in the last of these passages; and, since the publication of P 10 (1p), γαίῃ is there clearly the better attested reading. After its adoption, ἐπὶ γαίης disappears from the Iliad.

With a synonym χθών instead of γαῖα we have: κείται ἐπὶ χθονὶ πουλυβοτείρῃ Γ 195, κείται ἐπὶ χθονὶ Σ 461, ἐπὶ χθονὶ κείτο τανυσθείς Τ 483; τὴν μὲν κατέθηκεν ἐπὶ χθονὶ Ζ 473, τεύχεα κάλ' ἀποθέσθαι ἐπὶ χθονὶ πουλυβοτείρῃ Γ 89—and with substantially equivalent verbs ἐπὶ δὲ χθονὶ κάππεσον ἄμφω Ψ 731, ἔγχος μὲν κατέπηξεν ἐπὶ χθονὶ πουλυβοτείρῃ Ζ 213, νιφάδες . . . ἄς . . . κατέχευεν ἐπὶ χθονὶ πουλυβοτείρῃ Μ 158. There are two examples of the genitive: καὶ τοὺς μὲν κατέθηκεν ἐπὶ χθονός ἀσπαίροντας Γ 293, ἔγχος μὲν τόδε κείται ἐπὶ χθονός, οὐδέ τι φῶτα | λείσσω Τ 345.

The examples of ἐπὶ χθονός must, of course, stand before vowels. To some it may seem that metri gratia is an adequate explanation; but the problem is far from being so simple. The phrase stands before the bucolic dieresis, where, as Ahrens showed, hiatus is entirely unobjectionable. Then ἐπὶ χθονός may be no more than a half-intelligent 'correction' of the meter. It is to be compared with the more famous examples ἐλεγχείες, κακελεγχείες explained by Bechtel, *Lex.* 119-20, and by Cauer, *Grundfragen*³ 68-9.

The fact that one MS (Barroccianus 230, undated) reads ἐπὶ χθονὶ in Τ 345 adds an additional problem. Jachmann has recently drawn attention to υ 55, where a similar situation once existed.² There ἀφίκετο was the reading of Ω, but one MS U (Mon. Aug. 519^B, S. xiv) had ἀπίεστιχε with ἀφίκετο written above it. The isolated variant was given short shrift; but in 1906, P 19 (3a) was published, and it read

² Cf. *Class. phil.* 47.89 (1952).

ἀπέστιχε. Then Schwartz and Wilamowitz saw the merits of ἀπέστιχε, and it now stands in the text of Von der Mühl. Similarly the Barrocius manuscript may be the one to which the 'correction' of ἐπὶ χθονὶ did not permeate. Must we wait for a lucky discovery before recognizing its value?

3. A long plus passage Γ [396–418] was discussed in my *Athetized lines of the Iliad* 85–9. What was there said needs to be modified in two ways. (1) The absence of the lines from the edition of Zenodotus should have been assumed more confidently. The description on p. 40 of the way the epitomators of the scholia worked shows that if the statement of such a fact had been before them, they would have cut it out, if they were going to record the athetesis by Aristarchus. (2) Calling Christ's correction of Γ [416] 'necessary' was putting the case too strong.

The date of the passage is the first thing one should seek to determine. It was evidently composed to fill the place it now occupies, and consequently was made after 550 B.C. That it contains in [409] the only example of δούλη in the Iliad points further in the same direction, but still without determining how far we must go. Nor do we know how early in the 5th century the author of such a passage would have written in the Ionic alphabet. This effort leads to naught and we are forced to consider two possibilities.

Line [416] runs without important variant μέσσω δ' ἀμφοτέρων μητίσσομαι ἔχθεα λυγρά, and must be so printed in an edition that aims at the text of c. 150 B.C. The question whether that wording goes back to its author or to a μεταγραφάμενος remains open.

If the passage was first written in the Ionic alphabet, we have no reason to doubt that line [416] has come to us in its original form. We must seek to determine its meaning. Cauer, *Grundfr.*³ 109, says of the line that its words must mean: 'dass ich dir nicht von beiden Seiten Hass errege'. Similarly Mazon, *Introduction à l'Iliade* 274: '(si tu ne veux) que je te prépare au sein des deux peuples des haines sinistres'. English-speaking scholars have rendered it: 'and lest I devise grievous enmities between both, even betwixt Trojans and Achaeans' (Lang, Leaf, and Meyers); 'and lest I devise grievous hatred between both, Trojans alike and Danaans' (Murray in Loeb Library). Cauer goes on to declare: 'Der Ausdruck ist mindestens wunderlich.' I would say more positively that the phrase μέσσω ἀμφοτέρων can here mean only 'in the space between the two armies'. The phrase recurs ἐς μέσον ἀμφοτέρων συνίτην μεμαῶτε μάχεσθαι Z 120 = Υ 159³, and in μέσσω δ' ἀμφοτέρων σκῆπτρα σχέθον Η 277, where the meaning is 'in the space between the two men'.⁴ Compare ἐς μέσσον Τρώων καὶ Ἀχαιῶν ἐστιχόμενοι Γ 266 = 341. Then μέσον by itself means this space. Paris demands αὐτὰρ ἐμ' ἐν μέσσω καὶ ἀρηϊφίλον Μενελάον | συμβάλετε Γ 69, and his demand is restated Γ 90. Compare ἐς μέσσον ἰὼν Τρώων ἀνέργε φάλαγγας Γ 77, Η 55, ὥρσε δέ μιν κατὰ μέσσον Ε 8, ἀκόντισε . . . | ἀντικρὺ κατὰ μέσσον Π 285, καὶ δ' ἔθορ' ἐς μέσσον Δ 79, δινέουσι κατὰ μέσσον Δ 541, νεῖκος ὁμοῖον ἔμβαλε μέσσω Δ 444.

³ In Ψ, the μέσον is the arena, and the line had to be altered slightly: ἐς μέσον ἀμφοτέρω συνίτην μ. μ. Ψ 814.

⁴ The shift of meaning is slight, because each is the representative of his army.

For a long time this meaning of the phrase did not seem to make sense in this context: 'and lest I devise in the space between both armies grievous enmities [toward thee], and so thou perish in evil wise.' But very recently⁵ I have seen that it is not Aphrodite's devising, but the manifestation of the hatred inspired by her, that is meant to take place 'in the space between the two armies'. In other words we have here a slightly veiled threat that the goddess will cause Helen to be stoned to death in the space between the two armies.⁶

Now stoning to death is (perhaps) not mentioned in the *Iliad* elsewhere. There is a passage: ἀλλὰ μάλα Τρῶες δευδήμονες· ἥ τέ κεν ἦδη | λαΐνον ἔσσο χιτῶνα κακῶν ἐνεχ' ὅσσα ἔοργας Γ 56-7. That has been interpreted in three ways: 'you would have been stoned to death', or 'you would be lying in a stone grave', or a combination of these, cf. Studniczka, *Beiträge zur Geschichte der altgr. Tracht* 62₂₀ (Wien 1886). After mentioning that von Hartel has suggested to him that the passage makes reference to a 'Steinsarg oder Grabbügel', Studniczka says: 'Doch scheint eine Vereinigung dieses Gedankens mit der schon durch die Nachahmung bei Lykophron Alex. 333 bezeugten Grammatikererklärung nicht unmöglich: der Steinhaufe, welcher den Gesteinigten erdrückte, war zugleich sein Grabbügel.' The plus passage (if μέσσω be read) shows how its author understood Γ 56-7.

If the passage started in the Attic alphabet, this interpretation of it can still hold, but the ambiguity of the writing opens the way to another possibility, first seen by Christ and then approved by Cauer. Line [416] would have begun ΜΕΣΟΙ which can mean either μέσσω or μὴ σοί. Taking it one way or the other is no emendation. However, μὴ σοί must be followed by γ'; that is an emendation, and in my opinion an extremely easy one. In the Attic alphabet gamma and delta are enough alike to be easily confused; cf. the allegation, ap. Pausanias 7.26, that Γονέσσαν (B 573) was a Peisistratean blunder for Δονέσσαν. Apart from this, a μεταγραφάμενος who had taken ΜΕΣΟΙ as μέσσω would have seen the need of a connective, and would have had only δ' at his disposal.

Reading μὴ σοί γ' ἀμφοτέρων ... ἐχθεα gives a natural (not 'wunderlich') expression for what Cauer believed must be meant. In my *Ilias Atheniensium* I followed him and Christ. I should not do so now, for two reasons: (1) the chance that the plus passage was never in the Attic alphabet; (2) the unlikelihood that a blundering change of μὴ σοί γ' into μέσσω δ' should result in a text for which an easy, though not an obvious, interpretation can be given.

⁵ As a result of Chantraine's objection in a very kind review of my work, *Rev. de phil.* 26.56 (1952).

⁶ Lattimore translates: 'Lest I encompass you in hard hate, caught between both sides, | Danaans and Trojans alike, and you wretchedly perish.' That may indicate an understanding at least better than his predecessors had.

EPIGRAPHICA

JOSHUA WHATMOUGH

Harvard University

An index may or may not be a trustworthy shift for finding something a reader remembers having read without recalling just where he read it. No index is a substitute for reading. Accordingly the epigraphic items discussed in this paper were found by reading the entire contents of *CIL* 13 (in seven parts and two supplements, 1899–1943). The texts are selected for their linguistic interest.

(1) From Ripsdorf:

pro [salute] | impe(rii totiusque domus)
diuinae [] ancām Marti [|
Talliatium

In this votive inscription the object dedicated is described by the word *ancām*. Since the end of the line immediately preceding is broken, it is possible that a letter or two are lost. But the assumption is not necessary, nor is it likely. The only words ending in *-anca* to be found in the Latin dictionary are *branca*, *planca*, *zanca*. The last is said to be the name of a Parthian shoe; *planca* is 'plank', and *branca* is 'paw; branch'. None of these is appropriate. The sole objection to *ancām* is that it does not appear in the Latin thesaurus. But the objection is weak; either (a) *anca* is not Latin, or (b) it is to be added to the dictionary. There is nothing in its form that is decisive: meaning and etymology remain as a court of last appeal.

Greek has ἀγγος (τό) 'bowl, vessel', cf. ἀγγεῖον, OHG *ancha* 'testa' (Walde-Pokorny 1.38); with *k : g* we have in Umbrian *anšif* (acc. pl., *ID* 601) 'pateras' or the li'e (š from *k* before *i*); Gregory of Tours mirac. (de uirt. Iul.) 2.8 has *patenam et urceum, qui anax dicitur*; and finally *CIL* 2.416 (Lamas de Melido, Portugal) has *angom lamaticom* 'earthenware vase (?)', cf. *lama* (Ennius, Horace) 'soft earth, loam'. All this points to **ango-* 'vessel', in Germanic *anca-*, with *-m* (instead of *-n*) Latinized. The object of dedication is possibly the same as that described as *karvea* in Gaulish inscriptions.

(2) From Rigomagus (Remagen):

] Secundus | dec(urio) col(oniae)
Lug(dunensis) | ex euoc(atis) Aug(usti)
n(ostri) | cum pertic[a] | uiatoria
d(onum) d(edit) | u(otum) s(oluit) l(ibens)
m(erito)

Zangemeister in *CIL* proposed to read *Pertic[ia] Viatoria*, as proper names. This is contrary to all reason; *Viatoria* is a conceivable cognomen, but there is no gens *Perticia*. The difficulty lies in *cum*, which defines *donum* (and by extension *uotum*); cf. *CIL* 5.2790 *cum dono uotum soluit* etc. The *pertica uiatoria* is an emblem of the office of Secundus as *decurio*. But *pertica* is properly a measuring rod. It stands for an older **perc-ti-ca*, which is derived from the form that ap-

pears in Umbrian *perka-* 'rod, caduceus' and (with anaptyxis) Oscan *perek(a)* a measure of length (accompanied by a numeral), and perhaps in the Umbrian *perstu* (3 sg. impv.) 'fence off, delimit' and hence 'replace(?)' from **perk-sk-tōd*. It certainly is to be found in the Paelignian *o*-stem *[p]racom* 'enclosure' and in the Umbrian *ā*-stem *praco* (gen. pl.), which also means some enclosed locality in the lustral peregrination of the Atiedian brotherhood. Another derivative is *pergula*, *precula* 'crossbeam' and hence 'arbor, pergola'. The forms in *pr-* either show metathesis or, more probably, are alternants of the postulated **perek-*; as for *-rc-* : *-rg-*, compare *barca* : *barga*, *uerco-* : *uerga-breto*, *arcantodano* : *arganto-* (Lat. *argentum*), *Orcuarus* (Aquitania) : *Orgouarra*, in all of which the writing reflects the Brythonic *-rx-*.

But *pertica* came also to denote a measure of area as well as a linear measurement, and especially a lot assigned to a veteran soldier settled in a colony. In view of Neth. *perk* and Germ. *Pferch* it would seem that the dialect form as well as the Latin was known in the Rhineland. The Lex Ripuaria has *parricus*, and Italian *parco*, French *parc* point to a VL **parcum* with *a* from *e* before *r* followed by a consonant, cf. *nouarca* in App. Probi. This suggestion of the source of *parco* and related forms (Eng. *park*) was first advanced by Buecheler (*Umblica* 48) seventy years ago, but is hardly known to Romance linguists. To find *pertica* recorded at Remagen lends some plausibility to Buecheler's conjecture.

(3) Discussion centered around Oscan *ettuns* (Buck nos. 14–18²) is not yet stilled, though the etymology of the word is clear. It would have been in Latin **eitones*, **itones* nom. pl., presumably masculine beside the neuter *iter -inis*, and not very different from *iter* in meaning, cf. Greek masc. *ἄγων* properly 'a place of contest' beside fem. Latin *ambāgō*, *indāgō -inis*. That is to say, *ettuns* means 'broad streets, avenues' or the like which served as military routes from specified points inside the city to the city gates. All six inscriptions show the formula *eksuk amvianud ettuns*, i.e. 'ex hoc uico *itones', they all come from Pompeii, and all are of the same data, ca. 90 B.C., the period of the Social War, in the course of which Pompeii was taken by Sulla. A colony of his veterans was planted there, and they were constantly at feud with the previous inhabitants. This is a situation which repeats itself in the outposts of empire at a later date, when the Roman colonies had to face a hostile native population, especially along the frontiers. The Oscan inscriptions specify particular points such as numbered turrets, direction (e.g. a left turn), named gateways, and officers' quarters. Whoever is familiar with them, or will read Buck's translations, will recognize a similarity of content in the following inscriptions from the Rhineland. The chief difference is that the military route is placed under some divine protection, as was common with the superstitious imperial troops.

(a) From Colonia Agrippinensis (Köln):

Tutelae | pl[ateae?] M·M·L |
 sacrum | Ianuarius Incenus | emeritus
 u(otum) s(oluit) l(aetus) l(ibens) m(erito)

The expansion of M·M·L is debated. I conjecture that there is a reference to a fiftieth milestone, [ad] *milia quinquaginta* or *miliarium quinquagensimum* or the like. Another suggestion is to read *pl[atearum]* and to interpret the three litterae

singulares as each the initial of three local or other names in the neighborhood of Colonia Agrippinensis. What these may be is also conjectural. Many possibilities are given in *DAG* 221, 223, 224 (the local, divine, and personal names of Germania Inferior) e.g. *Lucretiae* (sc. *Lucretius uicus* at Colonia, Köln itself), *Mondiacensis* (not Mainz) out of many.

(b) From Kastell:

in h(onorem) d(omus) d(iuinae) | genium
pla(teae) p(ost) p(ortam) | pr(a)et(oriam)
L(ucius) L[] Vic|tor uet(eranus)
leg(ionis) | XXII d(edit) d(edicauit)

...

(c) Also from Kastell:

Iun(oni) reg(inae) | plat(eae) dext(rae)|
eun(tibus) Nid(am) | T. Veterius . . . fecerunt

(d) Also from Kastell:

I(oui) o(ptimo) m(aximo) | plat(eae) dext(rae)
e N(ida) | Adiutorius . . . fecerunt

(e) From Heddernheim:

in h d d | genium plateae noui
ui|ci cum edicula et ara | T Fl(avius)
Sanctinus mil(es) leg(ionis) | XXII

(f) Also from Heddernheim:

in h d d | genio platiae |
noui uici Aemi|lius . . .

(g) Also from Heddernheim:

in h d d | plat(eae) praetor(iae)
aram qui | et genium | Sattonius Gratus
d d . . .

The official in charge of a *platea* was called *platiodannus* (e.g. in *CIL* 13.6776 from Mainz, *platiodanni* [n. pl.] *noui uici* [gen. sg.], cf. Gaulish *dannus* 'magistrate', *arcantodan* [on coins] 'a monetary official'). There is also an inscription from North Africa, *CIL* 8.304 (Suppl. 11529), the opposite extreme of empire, which discloses exactly the same situation: *platea noua directa a porta militari*.

If any doubt remains concerning the meaning of *platea* (cf. Harsh in *Classical philology* 32.44-58 [1937], who took it as 'angiportus, uicus', i.e. 'lane, alley, street'), the following context and definitions will remove it:

Livy 25.25.8: *portis regione platearum patentibus stationes praesidiaque disposuit, ne quis in discursu militum impetus in castra fieri posset*. The scene is the siege of Syracuse in the Second Punic War: '(The walls of private houses were now like a fortification, protecting his camp.) At the gates which opened on to the streets [*portis regione platearum patentibus*], sentries and guards were posted, so that there might be no sudden attack on the camp while the soldiers were dispersed.' Isid. etym. 15.2.23: *plateae perpetuae ac latiores ciuitatium uiae sunt* (i.e. broad permanent city avenues); *CGL* 2 p. xii *plateae: uiae latae a porta in portam*. I conclude that *etruns* means 'thoroughfares'.

The materials discussed above, so far as they come from Germania Inferior and Superior, are noted in *DAG* note xlvi (with references to *CIL* 13 for all the epigraphic texts), where they illustrate non-Latin materials in the Latin inscriptions. The word *platea* is Greek, to be sure; but it found its way into Gothic (Mark 6.5 *plapjo*, for *-tjo* or *-þjo*?) as well as into western Germanic (German *Platz*). But the meaning in the Roman west under the empire is not 'place' (Italian *piazza*) but unequivocally 'street, avenue'.

THE GENITIVE SINGULAR OF *o*-STEMS IN GERMANIC

GUSTAV MUST

Cornell University

Except for the *o*-stem class, the formation of the IE genitive singular shows a clear and regular development from PIE *-os/-es* for all noun classes, in Germanic as well as in the other IE languages: Goth. *sunaus*, *gumins*, *fadrs*, *nahts*; ON *sonar*, *hana*, *fǫðor*, *nǣtr* (nom. *nǣtt*); OHG *fridō*, *hanen* or *hanin*, *fater*, *naht*; OS *sunō*, *gumen*; OE *sunā*, *guman*, *fæder*, *gēs* (nom. *gōs*); Lat. *hominis*, *patris*, *noctis*; Gk. *πατρός*, *κυρός*, etc. The gen. sing. of IE *o*-stems, however, is attested in forms whose explanation is problematical. This case is built in a number of markedly different ways in the IE languages: for Greek and Indic the indicated ending is the archaic *-sjo* (e.g. Hom. Gk. *ἱπποιο*, Skt. *aśvasya* from **ekyo-sjo*); for Illyrian Messapic it is *-ihi* (e.g. *blat̃ihi* 'Blossii'; cf. the gen. sing. of consonant stems in *-as*, e.g. *platoras*)¹; for Latin and Celtic it is *-i* (e.g. Lat. *servī*, Ir. *maicc* = older *maqi* from nom. *macc* 'son')²; while in Lithuanian, Lettish, and Slavic, the ablative singular has replaced the genitive (e.g. Lith. *diėvo*, Lett. *dieva*, Slav. *boga* 'of God'). In Germanic there is a form that has not yet been satisfactorily accounted for, namely Goth. *dagis*, ON *dags*, OE *dæges*, OS *dages*, OHG *tages*, etc. Because the IE languages differ so widely in the form of the gen. sing. of *o*-stems, most scholars tend to believe that no common proto-form for all the IE languages can be reconstructed.

This paper will discuss the problems presented by the form of the *o*-stem genitive singular in Germanic, and will attempt to find a way out of the difficulties that have hitherto confronted researchers.

As is evident from the examples given above, the formation of the *o*-stem genitive singular, in Germanic as in the other Indo-European languages, must be kept apart from the genitive singular of the other stems. The ending *-es/-os* results in Goth. *-s*, ON *-r*; in West Germanic it has completely disappeared. The ending of the *o*-stems (Gmc. *a*-stems)—Goth. *-is*, ON *-s*, OE *-es*, OS *-es*, OHG *-es*—is generally reconstructed as IE *eso/-oso*. This explanation, though widely held, has weaknesses that are obvious to every critical student: first, that an original intervocalic *s* in post-initial syllables yields Goth. *-s* when it has become final, but ON and WGmc. *-r*; second, that *e* in post-initial syllables becomes Gmc. *i*, as in OHG *lambir* from **lambizō* < **lombhesā*.

Numerous attempts have been made to overcome these difficulties. Brugmann suggested that the Gmc. genitive ending *-s* comes from IE *-e-so/-o-so*, transferred from the pronominal system (e.g. Goth. *þi-s*, *hwi-s*) to the nominal, with the *s* remaining voiceless because of the accentuation of the pronominal forms **þé-so*, **xʷé-so*.³ From a strictly formal point of view, this hypothesis would seem to be

¹ J. Whatmough, *The Osi of Tacitus—Germanic or Illyrian?*, *Harvard studies in classical philology* 42.143 (1931).

² F. Sommer, *Hb. d. lat. Laut- u. Formenlehre* 341 (Heidelberg, 1948).

³ Brugmann, *Gdr.* 2.2.162.

valid; but we must not overlook the facts which van Helten has already pointed out, viz. that the pronominal forms always point to PGmc. *z*, whereas the substantive forms point to *s*. If one wants to trace the *s* of the masculine and neuter pronominal genitive singular to IE *-eso/-oso*, it is difficult to understand why the *s* in this form should not have had the same treatment as in other pronominal forms.⁴ Hirt also held the view that it was hard to explain this *s* as coming only from pronominal stems.⁵

W. van Helten attempted to posit an ending *-esso* for Proto-Germanic, showing influence of the dative in *-esm* on the genitive.⁶ This theory is tempting, but has little chance of being right. Another theory is that the Gmc. accent was on the connecting vowel (*-ó-so*, *-é-so*).⁷ But this situation ought to have been reflected in the stem consonants: according to Verner's law, we ought to have forms like OHG **wolbes*, OS **wulbes* 'of the wolf'. The actual forms—*wolfes*, *wulfes*—show indisputably that the accent was on the stem vowel.

This is the present state of our knowledge. None of the theories mentioned above makes *-eso/-oso* acceptable as the origin of the Germanic form; and a glance at other Indo-European languages shows that in general there is scant evidence to support the theory of *-eso/-oso*. This suffix is required only for Greek and Old Church Slavic. It has been suggested that the gen. sg. ending *-osjo* has given Gk. *-oio*, but that *-oso* resulted in *-ov*. But as Brugmann showed in his Greek grammar, all the Greek forms can be derived from **-o-σio*.⁸ Schwyzler, too, finds an ending *-oso* unnecessary for Greek.⁹ There is an OCS pronominal genitive *česo* (n.) 'of which', a unique form that has been thought to retain an old ending *-so*. But it seems to me more plausible to derive the *-o* of this form from older *-os*; cf. *slovo* 'word' from **kleyos*, *kolo* 'wheel' from **kolos*.¹⁰

In short, there is no form at all whose explanation demands the positing of *-eso/-oso* as a genitive singular ending. Even Brugmann affirmed that the genitive ending *-so* appears definitely nowhere but in Germanic;¹¹ but as we have seen, this theory involves for the Germanic languages (except Gothic) so many phonological difficulties that even here it must be regarded as improbable and unusable. Forms hitherto explained as from *-so* must be explained otherwise.

Where else can the explanation be found? Only one solution seems possible to me: the Germanic form is to be explained by the well-known IE ending *-esjo/-osjo*. This ending, the origin of Gk. *-oio*, Skt. *-asya*, OIran. *-ahyā*, and Arm. *-oy*, has a wide geographic distribution. To associate the Gmc. and the IE endings has certain methodological advantages. First, the Gmc. form is now tied to a formation which is well attested in a number of IE languages, and no longer needs a special explanation. Second, this explanation involves no great phonological difficulties: Gmc. final *-s* comes from IE *-sjo* with the loss of final

⁴ Cf. W. v. Helten, *PBB* 34.105, *IF* 26.175.

⁵ *Hb. d. Urg.* 2.34.

⁶ *IF* 26.174.

⁷ E. Prokosch, *A compar. Gmc. grammar* 234 (Philadelphia, 1939).

⁸ Brugmann, *Griech. Gr.* 3 225 (Munich, 1900).

⁹ Schwyzler, *Griech. Gr.* 1.555 (Munich, 1939).

¹⁰ J. J. Mikkola, *Urslav. Gr.* 2.185 (Heidelberg, 1942).

¹¹ *Gdr.* 2.2.161.

-jo; cf. OE 2d Sg. imper. *dēm* from PIE **dhōmje*. It seems probable that a pre-consonantal *s* in a post-initial syllable could remain voiceless more easily than intervocalic *s*. Although we do not know exactly what influence IE *j* had on an immediately preceding *s*, we have parallel PGmc. cases which show that in a post-initial syllable *s* is preserved before *j*; thus *s* appears before the suffix *-jō* in certain old female designations, such as OHG *chebis(a)* 'concubine', MLG *keves kevese*, OE *cefes* 'pellex', OE *forleges* 'adulteress'.¹² The old *es/os*-stems are of particular interest for our problem, since in these the *s*-element belongs historically to the whole paradigm. The original inflexion of the *es/os*-stems has disappeared in Germanic, leaving only isolated traces, most of which can be exemplified in OE (Anglian and poetic), e.g. nom. acc. sg. *cælf*, *lomb*, *hréð*, gen. sg. *calfur*, *lombur*, dat. sg. *hróðor*, nom. pl. *calfur*, *lombur*, gen. pl. *calfra*, *lombra*, dat. pl. *lombrum*.¹³ In these examples the final *-z* of the nom. acc. sg. (from IE *-s*) has disappeared; the other cases have suffixes with *-r-* from *-z-* (also from IE *-s-*). The addition of a feminine *jō*-suffix to the neuter *es/os*-stems created a new feminine suffix *-isjō*. This suffix retains a voiceless *s*; but where the *s* of the *es/os*-stems is not supported by a following consonant, it nowhere preserves its voiceless character. The suffix *-isjō* became productive in Germanic and was also used for other stems, e.g. OHG *burissa*, OE *byres* 'borer' (to OHG *borōn*, OE *borian* 'to bore'), OS *lunisa*, OE *lynes* 'linchpin' (perhaps from an *es/os*-stem **luniz*), OHG *britissa* 'plank bed' (to OHG *bret*).¹⁴ It is suggested that these forms result from grammatical change (*-s- ~ -r-*), like OHG *langarra* 'deambulacra' and *chuburra* 'ratis'.¹⁵ Since the latter are restricted to OHG, we cannot be sure that they are really old formations. An alternation *s ~ z* has been assumed also for Goth. *agizi* 'axe' and *jukuzi* 'yoke'. But we ought not to neglect the Gothic law of dissimilation, according to which a voiceless consonant at the beginning of an unaccented syllable is followed in the succeeding syllable by *b*, *d*, *z*, and a voiced consonant by *p*, *t*, *s*.¹⁶ These Gothic examples are therefore not decisive in determining whether our examples originally had *s* or *z*.

The phonological development noted above—post-initial *s* remaining voiceless before *j*—strengthens the view that the voiceless *s* of the genitive singular in Gmc. *a*-stems was retained because it was followed in Proto-Germanic by *j*, the latter being lost in the later development of the Gmc. languages. We therefore conclude that the Gmc. ending *-s* in the genitive singular is related to the ending *-sjo* in the other IE languages.

This view is better able than any other to deal with the phonological difficulties which beset the explanation of this case form. We now see that the form is not isolated in Germanic, but is rather to be viewed in the framework of the IE family as a whole. The hypothesis strengthens our belief that *-sjo* was an original case ending of all *o*-stems, not only of the pronominal system.

Furthermore, the much discussed Latin genitive singular ending *-ī* can also

¹² F. Kluge, *Nominale Stammbildungslehre d. altgerm. Dialekte* 25 (Halle/Saale, 1926).

¹³ See Sievers-Brunner, *Abriss d. altengl. (ags.) Gr.* 42 (Halle, 1941).

¹⁴ Kluge, *Nomin. Stammbildungslehre* 45.

¹⁵ Kluge, *loc. cit.*

¹⁶ See W. Braune, *Got. Gr.* 50 (Halle/Saale, 1939).

be derived from PIE *-sjo*; for *-sj-* becomes *-j̥-* in Latin, as in *eius*, *quius* and *cuius* from **esjo(s)*, **q̥osjo(s)*.¹⁷ This explanation is improbable as long as we start from a PIE *-o-sjo* and try to connect Lat. *-ī* with Gk. *-οιο* from **-οοιο*.¹⁸ But when we consider that in Latin the ending *-is* (from *-es*) of the consonant and diphthong stems and the *ī-* and *ū-* stems is the ablaut form to *-os* (cf. Gk. *ποδός* : Lat. *pedis*), we can posit *-e-sjo* for the *o-* stems as well. Without phonological difficulties we derive the *-ī* through *-i-j̥(o)* from *-e-sjo*; for *e* becomes *i* before *s*, cf. 2d sg. pres. *agis* from PIE **aǵe-si*, Skt. *ājasi*. Before we can decide whether the Keltic formation in *-ī* is to be explained in the same way, we must wait for more exhaustive study of the Keltic *s*-clusters, in particular of *-sj-*.

One further difficulty is that the Gmc. languages do not show a uniform development of the vowel preceding the *-s*. The PN Runic inscriptions show *-as*, e.g. *goðagas*, *Asugisalas*. The oldest OE manuscripts likewise have *-æs* from *-as*, e.g. *dōmæs*, OWS *earmæs*; and we frequently find OS forms in *-as* as well (three times *-æs*), e.g. *armas*, *hobas*.¹⁹ These endings can be derived without difficulty from IE *-osjo*; they agree completely with Hom. Gk. *-οιο*. On the other hand, OHG always shows *-es*, e.g. *tages*, *wortes*, *lambes*, *fiskes*; and OS has *-es* beside *-as*, e.g. *hobes*, *grābes*. It has been suggested that this *-es* is an ablaut form of *-as*; but that is impossible, since non-initial *e* became Gmc. *i*. H. Naumann considers OHG *tages*, OS *dages* a compromise form between Gmc. *dagas* from IE *dhoghoso* and Gmc. *dagis* from IE *dhogheso*.²⁰ In like manner, R. Loewe suggested that the forms in *-as* coalesced with those in *-is* to an intermediate *-es*.²¹ In my opinion, OHG and OS *-es* can be better explained in a different way. It is to be noted that exactly the same form exists in the OHG *ja-* stems as in the *a-* stems, e.g. *hirtes*, *kunnes*. These *ja-* stems show *-es* as the regular development of the older *-jas* from PIE *-jo-sjo* in accord with the discussion above. From the *ja-* stems, then, this *es*-ending was taken over by the *a-* stems. The OS ending *-es* is a later form developed from *-as* through *-æs*, as in Old English.

The Goth. ending *-is* could be derived without difficulty, as far as Gothic is concerned, from *-eso*; but in order to unite Gothic with the other Gmc. languages we are justified in assuming that the Goth. *-is* comes from an ablaut variant of the ending just mentioned, viz. *-esjo*. Thus the same proto-form can be posited for Gothic as for Latin.

The Germanic pronominal genitive singular also ends in *-s*, as Goth. *hwis*, *þis*, *is*, ON *hves*, *þes(s)*, OE *hwæs*, *þæs*, *his*, OS *hwes*, *thes*, *is*, OHG (h) *wes*, *des*, neuter *es*. These forms could be plausibly derived from IE *-so*—if they were not pronouns, and if striking parallels did not exist in other IE languages. As already noted, pronouns do not usually undergo the same phonological developments as substantives; rather, since they are generally unaccented in the sentence, one would expect them to be treated like the unaccented syllables of substantives. Thus, Gmc. pronouns show an *r* from *z*, or a loss of *z*, in positions where the usual

¹⁷ Sommer, *Hb. d. lat. Laut- u. Formenlehre* 443.

¹⁸ Cf. Stolz-Schmalz, *Lat. Gr.* 268 (Munich, 1928).

¹⁹ See F. Holthausen, *Altsächs. Elementarb.* 91 (Heidelberg, 1921).

²⁰ *Ahd. Gr.* 28 (1922).

²¹ Loewe, *Germ. Sprachwissenschaft* 2.8.

sound correspondences would lead one to look for an *s*. Hence we should expect the PIE gen. sg. *k^weso* to yield PGmc. **x^wez(a)* and OHG **(h)wer*, just as the PIE nom. *k^wis* yields OHG *(h)wer* and PIE *k^wos* yields PGmc. *x^waz*, OE *hwa* (cf. also OHG *er* 'he' : Goth. *is*, Lat. *is*). A pronominal genitive singular like OHG *(h)wes* must therefore come from a different PIE form. As most of the IE languages indicate, the underlying PIE pronominal genitive singular ended in *-sjo*, e.g. Skt. *kásya*, *tásya*, *asyá* and *ásya*, Av. *čahyā*, Arm. *oroy* (to *or* 'which'), Hom. Gk. *τοιο*, Lat. *cuius* (< **k^wosjo-s*), *eius* (< **esjo-s*). The formation of the Germanic pronominal genitive singular in *-s* becomes clear when we relate it to the Indo-European formation in *-sjo*.

THE INSCRIPTION ON HELMET B OF NEGAU

KONSTANTIN REICHARDT

Yale University

Although modern scholars are becoming gradually less reluctant to regard the inscription on helmet B of the Negau trove as a solved riddle, I would propose a more cautious attitude. The only result to be expected from a new discussion of this enigmatic object is clarification, with a warning not to make use of its text for any far-reaching theories in historical, runological, mythological, or linguistic research.¹

Helmet B of Negau was found buried with twenty-five other bronze helmets in the year 1811 in Ženjak, Styria, not far from Negau, in the ancient border zone of Noricum and Pannonia Superior. The circumstances of the finding, with all related questions, have been thoroughly investigated by Reinecke in his article *Der Negauer Helmfund*. Twenty-one or possibly twenty-three of the helmets still exist. Seven show marks which appear to be more than mere scratches, and two carry full-fledged inscriptions in a North Etruscan alphabet (Reinecke 132-9). It was not until Marstrander wrote his brilliant article in *Symb. Osl.* (1925) that the inscribed helmets received a basically satisfactory interpretation. Helmet A carries the names of three successive owners with

¹ Citations, unless otherwise specified, refer to the following works. F. Altheim and E. Trautmann, *Vom Ursprung der Runen* (Frankfurt a. M., 1939). — F. Altheim and E. Trautmann-Nehring, *Kimbern und Runen* (Berlin, 1942). — H. Arntz, *Handbuch der Runenkunde*² (Halle, 1944). — H. Arntz, *Die einheimischen Runendenkmäler des Festlandes* (Leipzig, 1939). — O. Bremer, *Die Aussprache des R der urnordischen Inschriften*, *Festschrift tillägnad Hugo Pipping* 38-50 (Helsingfors, 1924). — K. Brugmann, *Grundriss der vergleichenden Grammatik der indogermanischen Sprachen*, Vol. 2, Part 2 (Strassburg, 1892). — S. Gutenbrunner, *Die germanischen Götternamen der antiken Inschriften* (Halle, 1936). — S. Gutenbrunner, *Historische Laut- und Formenlehre des Altisländischen* (Heidelberg, 1951). — H. Hirt, *Handbuch des Urgermanischen*, Vol. 1 (Heidelberg, 1931). — H. Jacobsohn, *Altgermanisches*, *ZfdA* 66.217-46 (1929). — M. H. Jellinek, *Geschichte der gotischen Sprache* (Berlin, 1926). — W. Krause, *Runeninschriften im älteren Futhark* (Halle, 1937). — W. Krause, *Was man in Runen ritzte*³ (Halle, 1942). — W. Krause, *Abriss der altwestnordischen Grammatik* (Halle, 1948). — P. Kretschmer, *Das älteste germanische Sprachdenkmal*, *ZfdA* 66.1-14 (1929). — W. Krogmann, *Germ. Harigasti*, *KZ* 64.269-71 (1937). — I. Lindquist, *HlewagastiR och Harigasti Teiua*, *Beiträge zur Runenkunde und nordischen Sprachwissenschaft* 86-102 (Leipzig, 1938). — R. Loewe, *Germanische Sprachwissenschaft*, Vols. 1 and 2 (Berlin, 1933). — C. J. S. Marstrander, *Les inscriptions des casques de Negau, Styrie*, *Symbolae Osloenses* 3.37-64 (Oslo, 1925). — C. J. S. Marstrander, *Remarques sur les inscriptions des casques en bronze de Negau et de Watsch* = *Avh. utg. av Det Norske Videnskaps-Akademi i Oslo*, 2. *Hist.-flos. Klasse*, 1926, No. 2 (Oslo, 1927). — G. Neckel, *Zur Inschrift des Helmes von Negau*, *KZ* 60.282-4 (1933). — J. Pokorný, *Indogermanisches etymologisches Wörterbuch*, Fasc. 7 (Bern, 1953). — E. Prokosch, *A comparative Germanic grammar* (Philadelphia, 1939). — P. Reinecke, *Der Negauer Helmfund*, 32. *Bericht der römisch-germanischen Kommission 1942* (Berlin, 1944, issued 1950). — E. Schwarz, *Goten, Nordgermanen, Angelsachsen* (Bern and München, 1951). — M. Schönfeld, *Wörterbuch der altgermanischen Personen- und Völkernamen* (Heidelberg, 1911). — F. Specht, *Germanisch Harigasti*, *KZ* 60.130-8 (1933). — W. Streitberg, *Urgermanische Grammatik* (Heidelberg, 1896). — J. Whatmough, *The Prae-Italic dialects of Italy*, Vol. 2 (Cambridge, Mass., 1933).

their patronymics: *siraku gurpi* | *isarni eisvi* | *tubni banuabi*. The first name probably shows a Scythian word in the nominative case—cf. Σιράκης, Σιραχος, and the tribe Σιρακες between the Maeotis and the Caspian Sea (Kretschmer, *Glotta* 30.186 f. [1943]); the name of his father is in the genitive case. The following four words can be explained as Keltic (Marstrander, *Symb. Osl.*). For the understanding of helmet B this inscription is of no significance.

Helmet B carries on the rim an inscription in retrograde order (right to left). Reinecke (139) gives an exact reproduction; a tracing of Reinecke's facsimile, natural size, is shown in Figure 1. A normalized version of the inscription is

HARIXASTITEIVA///IP

Only the letters preceding the three slant lines make sense; the last two letters, I P, are not explained.

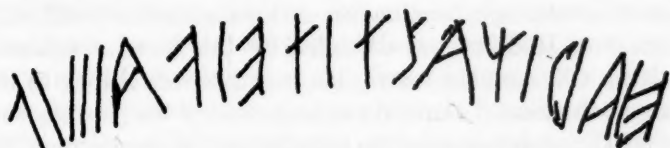


FIG. 1. INSCRIPTION ON HELMET B
Traced from Reinecke's Facsimile

Marstrander (*Symb. Osl.* and *AVAO*) broke ground by proving the Germanic character of *harigasti*; but his explanation of the second element *teiwa* as 'Teii Fa', 'Teii (filii) fa(ber)' or 'fa(brica)' was unfortunate. Kretschmer published the first article about Negau B as a wholly Germanic document under the proud title *Das älteste germanische Sprachdenkmal* (1929). No scholar after Kretschmer has expressed the opinion that the inscription is not Germanic.

The three elements of the inscription *hari*, *gasti*, *teiwa* can be connected easily with Germanic words: **harja-* (Goth. *harjis* etc.) 'host', **gasti-* (Goth. *gasts* etc.) 'guest', and **tiwa-* (ON *Týr* etc.) 'a god' or 'the god Týr'. The phonological condition of these nouns is thoroughly Germanic: IE *k* (**korjo-*, Mlr. *cuire*) appears as *h* (*hari-*), IE *gh* (**ghosti-*, Lat. *hostis*) as *g* (*gasti*), IE *d* (**deiyō-*, Skt. *dēva-h*) as *t* (*teiwa*), IE *st* (**ghosti-*) as *st* (*gasti*), and IE *o* (**ghosti-*) as *a* (*gasti*). These characteristics together establish the inscription as Germanic beyond all doubt.

Two peculiarities require and have received attention: the representation of IE stressed *ei* as *ei*, and the lack of a connecting vowel *-a-* in *harigasti*.

Even before the discovery that the Negau B inscription was Germanic, the view had been expressed that IE *ei* became *ī* in Germanic relatively late. The inscription of Xanten, *Alateiviae ex iussu i/ps(ius)/Divos medicus* (CIL 13.8606; Gutenbrunner, *Götternamen* 203), has the name of a Germanic goddess *Ala-teiv-*, with a second compositional element IE **deiy-*. Doubts were expressed (Hirt 1.38–9), but the Karelian *runko-teivas* 'rye-god' showed an unmistakable Germanic diphthong. Negau's *teiwa* is a valuable addition; the three examples together leave the impression that the documentation of IE *ei* as PGmc. *ei* only in the word 'god' cannot be accidental.

The form *harigasti* with syncopation of the connecting stem vowel is only slightly more difficult. I do not want to recapitulate all the known records, and shall confine myself to the statement that a syncopation of the *-a-* in *ja-*stems as first elements of compounds is relatively late in our documents. From ca. 100 A.D. we know *Chariovalda*, from the 3rd century *Ariogastus*, from the 4th century *Hariobaudes* (see the material collected by Schönfeld). In the large inventory of Germanic names known from Greek and Latin sources, *Charibertus* from the 6th century is the first example with syncopation. In runic inscriptions the earliest case of syncopation is the dative *kunimundiu* from the bracteate of Tjurkö, Sweden, ca. 550. But it is dangerous to draw hasty conclusions. The inscription of Birrens, Scotland, near Hadrian's rampart, has the clearly Germanic name of a goddess Harimella: *Deae Harimellae (sac)rum* (R. Much, *ZfdA* 36.44 ff., 63.19 ff.; Gutenbrunner, *Götternamen* 100; Krogmann 270). Although the exact chronological evaluation of this inscription is difficult, we know that the rampart of Hadrian was defended by the Romans against the Picts from 122 until ca. 400 A.D.; the inscription can therefore belong to the 4th century as well as to the second. An early syncopation of the *ja-*stem vowel is indicated also by OHG proper names of the type *Sibigell* (Gutenbrunner, *Götternamen* 15) and by developments in Gothic where medial *-ja-* became *-i-* in compounds after long vowel plus consonant or after two consonants; e.g. *arbi-numja*, *aglaiti-waurdein* (Jellinek 89).

This development is to be noticed not only in Germanic. I believe with Streitberg, *PBB* 14.170 (1889), that the relation between *kunimundiu* and Goth. *aljaleikops* is the same as between Lat. *mediterraneus*, *officiperda* and Gk. *μεσοπόνκτιον*.

Specht (133 ff.) tried to explain Negau's *hari-* as an IE *i-*stem **kori-s*, a variant of **korjo-s*. He drew his support from Lithuanian material, and Neckel (283) accepted the suggestion. I cannot find any reason for following Specht. Pokorny 615 lists only *koro-s*, *korjo-s* 'Krieg, Kriegsheer'; *korjo-no-s* 'Heerführer'. This is most probable.

I believe that the treatment of the Negau inscription in all its aspects has been deplorably superficial; the main purpose of this article is to stimulate more thorough investigation. My discussion is intended to be mainly linguistic, but I must touch on some nonlinguistic points as well, in order to clarify my attitude toward some of the important problems.

1. The owner of the Negau helmet, for whom the inscription was made, had Germanic blood—how much, it would be good to know. Let us not forget the Gothic bishop Ulfila, who was a mixture of Germanic and Cappadocian.

2. The person technically responsible for the inscription was most probably not of Germanic origin.

3. The method of communication between the owner and the alphabet-expert is unknown.

4. If statements 2 and 3 are correct, the skill of the inscriber as a phonetician or linguist becomes a significant factor.

5. If statement 4 is correct, the problem of possible 'mistakes' arises. But be-

cause of the lack of related material, it would be better to shelve the inscription than to interpret it with the help of assumed mistakes.

6. The place where the helmet was found proves nothing in regard to the place where it was inscribed.

7. For questions of chronology the archeological and epigraphic investigation has priority. The linguist cannot expect his conclusions to carry weight in the absence of comparable material.

The only thorough archeological investigation of the Negau helmets is that of Reinecke. He tells us that the Negau helmets are all different, though five main groups can be established. They represent the end-development of helmets known from Etruria in the 7th and 6th centuries B.C., developed after ca. 500 in central Italy. The Negau helmets are supposed to belong to the period of the early Roman emperors. Reinecke's treatise, an admirable contribution because of its general thoroughness, has one unfortunate characteristic: a longing for perfectionism. Reinecke admits the gaps in his archeological genealogy, but nevertheless tries to identify the burial of the Negau helmets with concrete political events. He reaches the conclusion that the helmets were hidden by a tradesman for their metal value, either during the expedition led by P. Silius Nerva in 16-15 B.C., or preferably during the fights in connection with the Illyrian-Pannonian revolt in 6-9 A.D. According to Reinecke, the owner of the Negau B helmet fought as a soldier in the Roman army. In other words, the man who owned the inscription *harigasti teiwa* died near the place where the twenty-six helmets were found.

Most of this strikes me as purely hypothetical. Reinecke does not admit that there is one very important piece of evidence against his conclusions: the alphabet on the Negau helmet. He comments briefly (176) on the chronology assumed for the North Etruscan inscriptions and demands its reconsideration. Before Reinecke, Marstrander and Messerschmidt and Kretschmer had all suggested earlier dates for the Negau inscription. Their arguments were based mainly on the shape of the letters, not of the helmet. Messerschmidt, in a very brief remark (523 f.), sets the time about 300 B.C. as the date of our inscription. Altheim-Trautmann, *Ursprung* 37, give 181 B.C. as the terminus ante quem and the 3rd century as most probable. In *Kimbern und Runen* 36 they set the date of the Negau inscription, without any supporting material, as ca. 200 B.C., but 'probably earlier'. For a linguist it is of no great importance to know whether the Negau inscription belongs to the 3rd century B.C. or to the beginnings of the 1st century A.D. It will become important only if additional records of the Negau type are found. But it is appalling to see that experts in North Italic archeology and epigraphy differ by more than two hundred years.

The alphabets used in North Etruscan inscriptions have been investigated thoroughly. Chronologically, one important feature in them is the influence of the Latin alphabet. Whatmough (520) assigns the inscriptions in which Latin influence is most evident to the early 1st century B.C. Altheim-Trautmann mention 90 B.C. as the year in which the Latin alphabet was victorious.

There are nine individual letters in the Negau B inscription. Among these the

forms of *h*, *s*, *t* suggest closeness to the Magrè alphabet. Whatmough (34, 520) assigns the Magrè inscriptions to the middle of the 3rd century. On the Negau helmet there is no influence of Latin letters. As long as the chronology of the North Etruscan inscriptions is not proved wrong, the Negau text cannot be considered later than the 2nd century B.C.

The words *harigasti* and *teiwa* are either two separate utterances or syntactically combined. In the first case *harigasti* should be the owner's name, *teiwa* the god's. If a syntactic relation is assumed, two interpretations are possible: as apposition or as subject-object.

Before we discuss the meaning of the inscription, we must determine which case forms are represented by the two words. Only after eliminating those case forms which contradict the results of historical Germanic linguistics can we undertake an interpretation of the text.

ACCUSATIVE. To take both words as accusatives is historically the simplest, in fact the only simple solution. IE **korǵoghostim deiŷom* would necessarily result in PGmc. **harǵ(a)gasti teiŷa*. Early North Germanic runic inscriptions show three examples of the *a*-stem accusative *staina*, and one not quite certain example of an *i*-stem accusative *hal(l)i* (Krause, *Runeninschriften* 242-3).

VOCATIVE. Special Germanic vocative forms are known only from Gothic in the declensions of the masculine *a*-, *i*-, and *u*-stems: *skalk*, *þiudan*; *juggalauð*; *sunu*, *daup̃u*, *Filippu*, *sunau*, *magau* (Jellinek 102, 105, 107). These forms presuppose the IE endings *-e*, *-i*, *-u*, *-ou*. Vocative forms in runic inscriptions cannot be made certain, although they are sometimes assumed. The best example is the inscription on the bracteate of Skodborg, North Slesvig, from ca. 500, in which according to some interpreters several vocatives appear: *auja alawin auja alawin auja alawin ja alawid*. Here *auja* may be a neuter noun 'good luck' (Goth. *awi-liups*, Skt. *āvati*), and *alawin*, *alawid* may be vocatives of the masculine proper names **Alawiniz*, **Alawiðaz*, known from OHG *Alwin*, *Alwid*. Krause, *Runeninschriften* 53, translates: 'Heil, Alawin! Heil Alawin! Heil Alawin und Alawid!' This is possible but far from certain. Marstrander, *NTS* 3.119 ff., took the text as Gothic and interpreted with good reason 'I protect Alawin (acc.) ... and Alawid (acc.)'.

Because IE *i*-stems form their vocative with *-i*, *-ei*, or *-oi* (Brugmann 542), *harigasti* could be, as Jacobsohn suggested (224 note 1), a vocative **korǵoghostei*. The difference in the development of IE *ei* in the root vowel of *teiwa* and the ending of *-gasti* could easily be explained by the difference in stress: stressed *ei* remained, whereas in unstressed position it had become a monophthong by the period of the Negau inscription.

Though we may accept *harigasti* as a vocative, we must reject *teiwa*. An IE *o*-stem would require the ending *-e*, zero in Germanic. To suggest early analogical influence of the nominative [**teiŷa(z)*] would be awkward after positing an archaic form for the *i*-stems.

GENITIVE. Marstrander (*Symb. Osl.* 59) suggested the explanation of *harigasti* as a Celtic genitive of a Germanic name. This opinion is acceptable, but he did not succeed in offering a satisfactory explanation of *teiwa*.

It is not impossible to explain *harigasti* as a correct Germanic genitive. The IE *i*-stems had the endings *-eis* or *-ois*. Germanic feminine *i*-stems such as Goth. *anstais*, ON *nauðar* have IE *-ois*. The IE *u*-stems had the corresponding endings *-eus*, *-ous*; the latter is known from Goth. *sunaus*, ON *sonar*, OE *suna*, OS *suno*, OHG *fridō*. The Germanic masculine *i*-stems prefer analogical genitive forms after the *a*-stems (Goth. *gastis* etc.); but the runic inscription on the stone of Årstad, Norway, from the 6th century, has preserved a genitive *-winar* 'friend's', and Old Norse has *staðar*, both presupposing an earlier *-aiz*, IE *-ois*. Umbrian *Herentaleis* and Oscan *punes* presuppose IE *-eis* (Brugmann 573-4); and if we take into account that Indic, Baltic, and Slavic *i*-stem genitives may be explained from either *-ois* or *-eis*, *harigasti* also can be derived from IE **korjo-ghosteis*, PGmc. **hariagastiz*. The problem of IE *-s* : PGmc. *-z* : Negau zero will be discussed under the nominative.

teiwa cannot be a genitive. All Germanic languages preserve IE *-s* in the genitive of *a*-stems.

DATIVE. The formation of the dative of Germanic masculine *i*-stems is not entirely clear. Goth. *gasta*, OS OHG *gaste* indicate analogy after the *a*-stems (Streitberg 242). Old Norse shows the vast majority of nouns without ending and with *i*-umlaut in its proper place, e.g. *gest*, *stað*. Prokosch (245) should not have used the paradigm form *geste*. The ON forms can be compared to early West Germanic forms such as OS *hugi* (Codex Monacensis), OE *daeli* (Epinal Glosses), OHG *quidi*; they can all result from an IE instrumental in *-i* (Skt. *matī*). Early runic forms are not available, but we should expect **gasti*, *staði* (Krause, *Abriss* 61; Gutenbrunner, *Laut- und Formenlehre* 95). It is incomprehensible how Schwarz (77) can posit a Gotho-Norse **gastai*. I agree with Marstrander, *Symb. Osl.* 64: 'Si l'on suppose que la forme grammaticale de l'inscription est germanique, *Harigasti* peut également s'expliquer comme un datif singulier, identique au datif v.-nor. *gest* (issu de **gasti*, *gastē-i*).'

Kretschmer (6) was first to state that both *harigasti* and *teiwa* could be understood as dative forms, but he failed to offer any specific linguistic comment. Neckel (283-4) used Kretschmer's suggestion for his interpretation 'to the god Harigast', which has received wide acclaim. But he did not consider it necessary to justify the ending *-a*.

To be sure, Gothic *a*-stems have the dative ending *-a*. This can be derived from various IE endings: locative in *-oi*, instrumental in *-ē*, *-ō*, ablative in *-ēd*, *-ōd* (Brugmann 617; Prokosch 235). The other Germanic languages have *-e*: ON OS OHG *arme*, OE *earme*. This ending presupposes PGmc. *-ai*, IE *-oi*. The earliest certain runic form is *woduride* on the stone of Tune, Norway, from ca. 400. Krause's suggestion (*Arkiv f. nord. fil.* 48.156 ff. [1932] and *Runeninschriften* 150-1) that the stone of Möjebro, Sweden, also from ca. 400, should have the inscription *frawaradaR ana hahai slaginaR* 'Frawarad on the horse slain' would imply the earlier ending *-ai*, but the older reading *ana haha is slaginaR* is preferable. Cf. O. v. Friesen, *Upplands runstenar* 4 (Uppsala, 1913).

The Gothic dative ending *-a* is a peculiarity without parallels in North or West Germanic. Kretschmer (5) states that it is 'certainly possible' to assume for 'an old West Germanic dialect' the same ending as in Gothic. I do not follow this.

A Proto-West-Germanic dialect is here supposed to have had an ending which is unknown in later West Germanic, and this assumed ending is supposed to have developed between the 3rd century B.C. and 10 A.D., producing a result known from the distant Gothic in the late 4th century A.D. But even if we had faith in such peculiar developments in general, we still could not explain this Proto-West-Germanic *-a* historically. IE *-oi* resulted in WGmc. *-e*; IE *-ē*, *-ēd* also in WGmc. *-e*; and IE *-ō*, *-ōd* in WGmc. *-u*. There is no place in West Germanic for an *a*-stem dative in *-a*.

For those who insist that both *harigasti* and *teiwa* are datives and that the latter has a legitimate ending, it is of course simplest to assume a mistake in the inscription. Arntz (*Handbuch*²) simplifies the matter by printing *harigasti teiwai* opposite the photograph of the helmet on Plate 1. Schwarz says (72-3): 'Der Dat. Sg. ging auf *-ai* aus, wohl noch in den ältesten an. Runeninschriften und in *Harigasti teiwa* "dem Gotte Harigast" des Helmes von Negau, wenn man *teiwai* für das inschriftliche *teiwa* einsetzen darf.' Nein, das darf man eben nicht! (Incidentally, the same author [49] states that *harigasti* corresponds to IE *Corioghosti*.)

NOMINATIVE. The suggestion that *harigasti teiwa* might be nominative was first made by Neckel (Kretschmer 6), but he later abandoned this view in favor of datives. Lindquist, who holds that we should assume two accusatives, at the end of his article (102) mentions nominatives as his second choice.

IE *-s* as a masculine nominative ending is preserved in Gothic (*s* > *z* > *s*), appears in North Germanic runic inscriptions as *-R*, and is lost in West Germanic. The date of the WGmc. development is unknown. Bremer (42-3) expressed the opinion that the disappearance of *-z* in West Germanic could be established for a very early time because of the proper names *Nasua* (Caesar, BG 1.37), *Catualda* (Tacitus, Ann. 2.62), and *Chariovalda* (ib. 2.11); cf. Schönfeld s.vv. *Nasua*, which has not been satisfactorily explained, must be left out of account; but *Catualda* and *Chariovalda* are undoubtedly Germanic names, with a Keltized form of the first element: *Catu-* for Gmc. **Haðu-*. Bremer may be right, or at least cannot be proved wrong; but the ending *-a* may also represent IE *-ón*, especially since compounds with *-vald-* as second element show a mixture of strong and weak declensions. The Old Norse poetic vocabulary provides *allvaldr* : *allvaldi*, *einvaldr* : *-valdi*, *folkvaldr* : *-valdi*. Unfortunately, the situation of the *n*-stem nominatives in all three genders is still badly confused in Germanic; a definitive statement could be made only on the basis of unwarranted confidence in one or another of the current theories.

It is possible that a PGmc. **harigastiz teiyaz* had become *harigasti teiwa* as early as the time of the Negau inscription; if so, the inscription definitely shows West Germanic features. On the other hand, we may ask how **harigastiz teiwaz* would have been expressed in a North Etruscan alphabet. Whatmough's table (502) shows that the sound *z* is represented by a special letter in the alphabet of Sondrio; but this is not the alphabet used on the Negau helmet. The alphabets of Magrè and Este, which are closely related to the Negau letters, have no certain representation for *z*. The crucial question is probably unanswerable: how would a user of the Magrè or the Este alphabet have written a Germanic *-z*?

Jacobsohn (224 note 1) makes an interesting comment which I should like to quote, because it is easily overlooked. After remarking that some of the North Etruscan inscriptions show the influence of the Etruscan language, he adds: 'nun ist es fürs etruskische charakteristisch, dass das sog. nominativ -s nur an gentilicia antritt, wenn sie mit einem pränamen verbunden sind. es fehlt bei pränomina und alleinstehenden gentilicia. darf man annehmen, dass *Harigasti* mit seinem fehlenden -s eine etruskierte germanische form darstellt?' I am not competent to pass judgment; but Specht, who was certainly competent, called Jacobsohn's idea 'ansprechend' (130).

The result of our survey is that *harigasti* can be more or less plausibly derived from five IE case forms: nom. **ghostis*, gen. **ghosteis*, instr. (Gmc. dat.) **ghosti*, acc. **ghostim*, or voc. **ghostei*, while *teiwa* can come only from the nom. **deiyos* or the acc. **deiyom*. In my opinion this result suggests five possible interpretations:

(1) Accusative plus accusative: '(X invokes) Harigastiz the god' or '(X invokes) Teiwaz Harigastiz'. The slight but syntactically important difference between these two translations depends on whether we consider *teiwa* in apposition or *harigasti* a bahuvrihi compound.

(2) Nominative plus accusative: 'Harigastiz (invokes) Teiwaz'.

(3) Genitive plus nominative: '(Helmet of) Harigastiz. (God) Teiwaz'.

(4) Genitive plus accusative: '(Helmet of) Harigastiz. (He invokes) Teiwaz'.

(5) Nominative plus nominative: 'Harigastiz (the owner). Teiwaz (the god)'.

Neckel's interpretation 'to the god Harigast' cannot be correct, though it has been often repeated. Arntz and Schwarz accept Neckel's dative theory. Krause (*Runen* 12) quotes his translation with some hypothetical remarks of his own. Pokorny (615) is a little more cautious: 'PN altgerm. *Hari-gasti* (Dat.), vielleicht Göttername'. Neckel's verdict was based not on linguistic evidence, but on his very positive opinions concerning the ancient Germanic civilization. In the discussion of the Wotan-Óðinn cult in Germanic religion he took a strong stand in favor of the early pan-Germanic significance of this god, and tried hard to identify him with the *regnator omnium deus* in Chapter 39 of Tacitus' *Germania*. It was therefore a matter of considerable importance to Neckel to find Wotan mentioned in the earliest Germanic inscription.

In Old Norse sources Óðinn appears quite frequently as a guest; in the *Hervarsaga* he calls himself *Gestumblindi*, probably 'the one-eyed guest'. His close relation to war and warriors is well established. In Old Norse poetic tradition he has more than a hundred cognomens, and no one familiar with Eddic and Skaldic poetry would be surprised at seeing him referred to as *Hergestr*. Neckel believed, often with very good reason, that the Old Norse literary material represents a very old Germanic tradition: so far as Germanic customs and the Germanic philosophy of life are concerned, Tacitus' *Germania* has its best commentary in the Icelandic saga.

But here we are dealing with just two words—two words without a context; and we must not forget that there is a gap of thirteen hundred years between the Negau inscription and the Old Norse manuscripts. Especially important is the fact that *Harigastiz* as a name is not isolated: a 10th-century OHG source

mentions the proper name *Herigast* (Marstrander, *Symb. Osl.* 59), and we have a synonym in the Franconian *Volcast*, probably *Folk-gast* (Lindquist 101).

For *teiwa* Old Norse literature permits two interpretations. The word may refer to the god Týr. In Norse mythology he lives in the shadow; but he is the hero of a myth which tells of the binding of the Fenris-wolf and explains the god's loss of one of his hands. His great importance in early Germanic religion is reflected by his appearance in the name of the second day of the week, and by the fact that the *t* rune is named for him. Of the twenty-four rune names only two can be safely connected with the names of individual deities: **Tīwaz* (*t*) and **Ingwaz* (*ng*); see Reichardt, *Runenkunde* 30-42 (Jena, 1936).

Old Norse records, however, contain also an appellative *týr* 'god'. This is common in compounds, such as *Gautatýr* 'god of the Gautar', and appears in poetry seven times in the plural form *tívar*. The only example of the singular simplex *týr* occurs in a poem of the 9th century; consult F. Jónsson, *Lexicon poeticum* 576 (1931). A combination *Hergestr týr* 'the god Hergestr' would not surprise us in Old Norse, but there is no example of the appellative *týr* following a name in apposition.

Neckel's interpretation of the Negau text is linguistically unsound but mythologically fascinating. If he had called the words accusatives or nominatives instead of datives, he would have been linguistically on firmer ground. Lindquist, in his very stimulating paper, chose the unassailable accusative theory. He did not state what verb he thought of as underlying the text, but likely candidates can easily be supplied. Among the verbs that mean 'to invoke', 'to pray to' there are several with a predilection for accusative objects, e.g. Goth. *anahaitan*, ON *nefna*, OE *weorðian*, *herian*, (ge)biddan, *eaðmēdan*.

The main purpose of Lindquist's article was to show that *harigastiz* is a bahuvrihi compound; he translated it 'som har en här till gäster', i.e. 'who has a host as guests'. There is no room here to discuss Lindquist's general theory concerning Germanic names in *-*gastiz*, but as it applies to the Negau inscription it must be called wrong, since it contradicts a rule of Old Germanic syntax. Before the influence of Latin *dominus* and the like changed the original pattern, the title always followed the name; ON *Haraldr konungr* and early OHG *Christe druhtine* (in Isidor) are examples of the earlier usage, and many others occur in the Anglo-Saxon Chronicle. This rule was established in several exciting sessions of Neckel's seminar in the early twenties in the University of Berlin; Neckel published a brief summary of the results in his lucid article on Germanic syntax, *Acta phil. scand.* 1.8-10 (1926-27). If we assume that the two words on the Negau helmet are in apposition, the only possible interpretation is 'Harigast the god'.

The Negau inscription is unique. A-priori expectations concerning its meaning are idle. For mainly chronological reasons, Reinecke (172-6) insists that the text must contain the owner's name. He acknowledges (173) that votive inscriptions on helmets are known from earlier times, but denies their existence during the 'late-Republican-Julian-Claudian' era. Since he is convinced that the Negau inscription is not earlier than the latter part of the 1st century B.C., he refuses to see in its text an exception to this principle.

I must confess that this reasoning confuses me. If for the moment we accept Reinecke's chronology, we shall still find his additional arguments unconvincing. What do we know about the psychological background of the Negau text? The Germanic owner may have been a common soldier, or he may have been an officer in a Roman or some other military unit. He may have been a polyglot and even a skilful user of the alphabet, or he may have been illiterate. He may have seen other inscribed helmets, and may even have watched an expert inscribing one. He may have been following a tradition, though in that case the question remains: which one? He may have known that other soldiers dedicated their helmets to deities in the dative case, or he may have known that they had their own names inscribed in the genitive. All these are possibilities, but none of them is helpful. We shall never know what our Germanic warrior wanted to express. Even if Reinecke's main thesis should prove correct, we must reject every a-priori argument concerning the sense of the Negau inscription.

One of the most important results of modern runological studies is our belief that the runic alphabet was developed in close connection with the North Etruscan alphabets, perhaps during the 2nd or 1st century B.C. The Negau helmet gives no help to runology; but we may ask whether the psychology of the early runic inscriptions may not help us to interpret the Negau text.

Arntz (*Runendenkmäler* 468) lists continental runic inscriptions which in his opinion have a magical or votive meaning. These are of no significance for our problem. The inscription on the fibula of Kärlich in western Germany, *wodana heilag* or *wodani hailag*, is a falsification in spite of Arntz's hesitant discussion (266-73). For our purposes, the most important of the continental inscriptions is the one on the clasp of Nordendorf in south Germany, dating from ca. 600: *awaleubwini wigiponar wodan logapore*. *Awa* is a feminine, *Leubwini* a masculine proper name, denoting either the owners or the donors of the object. The compound *wigiponar* has as its second member the name of the god OHG *Donar*, ON *þórr*; it can be considered a variant of ON *Vingþórr*. *Wodan* is OHG *Wotan*. The last word, *logapore*, is not yet explained, but can be connected with the ON god *Lóðurr* (consult Krause, *Runeninschriften* 202-6; Arntz, *Runendenkmäler* 277-300). There is no doubt that the Nordendorf clasp has a votive inscription, but what case form did the inscriber choose for the names of the gods? Both *-ponar* and *wodan* can be nominative, vocative, or accusative.

Among North Germanic runic inscriptions of the earlier period, the bracteate of Arum, Sweden, from ca. 550, perhaps mentions a deity in the dative: *ehe ik akaz fahi* 'to the horse I Akar paint (these runes)'. The first word is adequately explained as PGmc. **ehyai*, Goth. *aihwai* (cf. *aihwatundi*); the significance of the horse in Germanic religion is well known; Óðinn is the horse-rider, and bracteates have a predilection for pictures of horse and rider. Earlier than all bracteates is the alphabet inscription of Kylver, Gotland, ca. 400, where a separately inscribed word *sueus* baffled runologists until Marstrander explained it as a magical writing for *eus* 'horse' (written in both directions, from the central *e*; see NTS 3.67 ff.). If this is correct, the noun is in the nominative.

One of the most surprising features of the Negau text is the relation of its *teiva* and the mention of the same god in runic inscriptions. There cannot be

any real connection between these inscriptions and the helmet; yet I am unable to subdue the impression that there is a psychological link of some sort. The Eddic poem *Sigrdrífumál* advises the warrior, in order to insure victory, to carve the *t* rune on his sword, and to mention the god's name twice (*nefna tysvar Tý*). Separate *t* runes are known from North Germanic inscriptions and are interpreted in the light of the Eddic verse. The stone of Kylver, mentioned above because of its *eus*, has at the end of the complete alphabet inscription a tree-like symbol that looks like a multiple *t* rune. Bracteates from Allensø in Denmark and Dannenberg in Hanover combine *t* and *e* runes; Krause has explained these as symbolizing the god Týr and the horse deity (*Runeninschriften* 38-40). Two fibulas from the Viking period, found in Gotland, have a *t* preceding the name of the rune-master and *siktiir* after the owner's name. The latter word is identical with ON *sigtýr* (nom.) 'victory god' (Lindquist 100). The runic material is neither rich nor conclusive; but at least it shows that the mention of *teiwa* on the Negau helmet is not unique.

I do not think that any conclusion can be reached that will not be open to grave doubts. For myself, I believe that the Negau inscription contains two nominative forms, whose ending *-z* is absent either because of an early West Germanic development or because of some peculiarity of the North Etruscan alphabet used for the inscription. I cannot disprove the translation 'Harigastiz the god', but I prefer to see *Harigasti* as the name of the owner, and *Teiwa* as that of the god.

For historians the helmet of Negau would have great significance, if only its date could be established. For runologists it provides the puzzling appearance of a pre-runic *t*. Mythologists will be delighted to find Wotan, or perhaps only Týr, in such an early record. Linguists will have to content themselves for the present with the knowledge that *harigasti teiwa* is the oldest Germanic text.

THE OATHS OF STRASSBURG: PHONEMICS AND CLASSIFICATION

ROBERT A. HALL JR.

Cornell University

It has been customary in Romance philology to study and analyze texts in terms of their orthographical tradition, even where—as in the case of the famous Oaths of Strassburg of 842—their phonology has been the subject of intensive discussion and interpretation. At the present stage of linguistics, however, this is not sufficient; since every written document represents an act of speech, it is our job to determine as closely as possible exactly what we assume that act of speech to have been, and to analyze its systematic aspects just as we would those of living informants, as far as our evidence permits, making use of historical and comparative linguistic data for synchronic interpretation. This analysis we should then present in consistent structural terms, to make it most usable in further historical and comparative work. It is our intention to do so here for the Oaths of Strassburg, presenting (1) each of the Oaths in its traditional orthography, with a translation; (2) a discussion of the phonemic system that we assume for the Oaths on the basis of descriptive and historical criteria; (3) a tentative phonemic transcription of the Oaths; and (4) a discussion of the implications of this re-interpretation for our classification of the language of the Oaths.

1. TRADITIONAL ORTHOGRAPHY AND TRANSLATION. We reproduce the text of the Oaths as it is accepted by virtually all editors. Of the emendations customarily made, we include those of *dist* to *dift* in the first Oath, and of *sue* to *sua* in the second. On the other hand, we do not follow the earlier custom of emending *n lostanit* in the second Oath to *lo suon franit*, since Tabachovitz (*Etude*, ch. 9)¹

¹ References are to the following works, cited by authors' names alone; abbreviations in this list are explained in the author's *Bibliography of Italian linguistics* (Baltimore, 1941).

G. Baist, *Alfranzösisch dh (ð) in altenglischen und altdeutschen Lehnworten* [with F. Kluge], *ZRPh.* 20.322-34 (1896).

G. De Poerck, *La diphtongaison de ϵ et de φ en ancien français et la palatalisation de \bar{u}* , *Handelingen van het Negentiende Vlaamse Filologencongres* 141-7 (Brussels, 1951).

A. Ewert, *The Strassburg Oaths*, *TPS* 1935.16-35.

R. A. Hall Jr., *The reconstruction of Proto-Romance*, *Lg.* 26.6-25 (1950).

P. Marchot, *Petite phonétique du français pré littéraire* (Fribourg, 1901).

J. Melander, review of Tabachowitz *Etude* [below], *StN* 6.164-70 (1934).

W. Meyer-Lübke, *Die Strassburger Eide und die vokalischen Auslautgesetze*, *ZRPh.* 12.526-7 (1888).

R. Politzer, *Free tonic closed \bar{e} in the Oaths of Strassburg*, *RPh.* 5.318-20 (1952).

W. Suchier, review of Tabachowitz *Etude* [below], *LBl.* 59.265-6 (1938).

A. Tabachovitz, *Etude sur la langue de la version française des Serments de Strasbourg* (Uppsala, 1932).

A. Tabachovitz, *Quelques remarques complémentaires sur la langue des serments français* (Härnösand, 1936).

W. von Wartburg, review of Ewert [above] and Tabachowitz *Quelques remarques* [above], *ZRPh.* 57.655-6 (1937).

A. Wallensköld, *Les serments de Strasbourg*, *Philologische Schriften . . . Karl Voretzsch dargebracht* 87-104 (1927).

The foregoing list does not pretend to be a complete enumeration of all the discussions of the Oaths of Strassburg. For a bibliography to 1932, see Ewert 32-5.

has shown that the original text may be interpreted as *non lo s(e) tanit* 'does not keep it'.

1.1. OATH OF LOUIS THE GERMAN. *Pro deo amur & pro christian poblo & nostro commun salvament, dist di in avant, in quant deus savir & podir me dunat, si salvarai eo cist meon fradre Karlo & in aiudha & in cadhuna cosa si cum om per dreit son fradra salvar dift, in o quid il mi altresi fazet & ab Ludher nul plaid nunquam prindrai qui meon vol cist meon fradre Karle in damno sit.*

'For the love of God and for the Christian people and our common well-being, from this day onward, insofar as God grants me to know how and to be able, I will help this brother of mine Charles, both in aid and in every matter, as a man should help his brother, insofar as he does likewise by me, and with Lothair I will make no agreement which by my will might be harmful to this brother of mine Charles.'

1.2. OATH OF CHARLES THE BALD'S SOLDIERS. *Si Lodhuuigs sacrament, que son fradre Karlo jurat, conservat, et Karlus meus sendra de sua part non lo s tanit, si io returnar non lint pois, ne io ne neuls, cui eo returnar int pois, in nulla aiudha contra Lodhuuig nun li iv er.*

'If Louis keeps the oath which he swore to his brother Charles, and Charles my lord for his part does not keep it, if I cannot deter him therefrom, neither I nor anyone whom I can deter therefrom will be of any help to him in this matter against Louis.'

2. PHONEMIC SYSTEM. In this section, we list the individual phonemes we assume for the dialect of the Oaths, with their orthographical representation, their presumed allophones, and examples (with phonemic transcription, orthography, phonetic transcription, and gloss for each).

2.1. VOWELS: /i e a o u ə/, and a contrast of tenseness vs. laxness applying only to the stressed mid-vowels /e o/, which we symbolize by writing /^ː/ after tense vowels and leaving lax vowels unmarked.

/i/ i: [j] at the beginning of a syllable before a syllabic, [i] at the end of a syllable after a syllabic, [i] as syllabic: /kristián/ *christian* [kris'tjan] 'Christian'; /ió/ io [jo] 'I'; /pláit/ *plaid* ['plajt] 'agreement'; /dréait/ *dreit* ['drejt] 'right'; /dí/ di ['di] 'day'.

/u/ u: [u] as syllabic (presumably also [w] and [ʊ] parallel to [j] and [i]): /núl/ *nul* ['nul] 'no (adj.)'.

/e/ e: [e]: /déus/ *deus* ['deəs] 'God'; /konsérvat/ *conservat* [kon'servat] 'keeps'.

/o/ o: [o]: /óm/ *om* ['om] 'man'; /kósa/ *cosa* ['koza] 'thing'.

/a/ a: [a]: /salvár/ *salvar* [sal'var] 'to help'.

/e/ e before i, otherwise i: [e]: /dréait/ *dreit* ['drejt] 'right'; /savéar/ *savir* [sa'ver] 'to know'; /céast/ *cist* ['t'est] 'this' (m. obl. sg.); also, morphophonemically, in /prendrái/ *prindrai* [pren'draj] 'I shall take'.

/o/ u: [o]: /kóam/ *cum* ['kom] 'as'; /amóar/ *amur* [a'mor] 'love'; /nóan/ *nun* ['non] 'not' (stressed); also, morphophonemically, in /rətorjár/ *returnar* [rətor'nar] 'to turn back'.

/ə/ e, o, a, u: [ə]: /frádrə/ *fradre, fradra* ['fraðrə] 'brother'; /kárle/ *Karlo, Karle* ['karlə] 'Charles' (obl.); /kárles/ *Karlus* ['karləs] 'Charles' (nom.).

Of the eight vowels that we postulate, the first five are assumed by all commentators. With regard to the other three there is no consensus of opinion. The

interpretation of *i* and *u* (< 'Vulgar Latin' or 'Proto-Italo-Western Romance' /é/ and /ó/) has been the subject of considerable debate. There are two main interpretations current: that by which they are taken to stand for the diphthongs /éi/ and /óu/ respectively (thus most commentators, e.g. Wallensköld, Melander, von Wartburg, Politzer), and that by which they are assumed to represent simple vowels (e.g. Marchot, Tabachovitz *Etude*, De Poerck). Yet there is no evidence to connect the Oaths with those regions of southern Gallo-Romance where /é/ > /i/ and /ó/ > /u/, and De Poerck's effort to prove (145) that *i* and *u* in the Oaths stand for 'short' [i] and [u] shows confusion between the phonetic and the phonemic levels of analysis, since what he sets up as 'short *i*' and 'short *u*' must have had the same phonemic function as the tense /é/ and /ó/ of other analysts, regardless of their phonetic identity. As for the supposed diphthongization, there is no internal evidence at all to point to such a phenomenon in the Oaths, and there is direct internal evidence to the contrary, in the use of *ei* in *dreit*, which must stand for /dréait/ < /diréaktu/ (as pointed out long ago by Baist 330). If the scribe had had /éi/ in */savéair/, he would have used the same graph *ei* for writing this word as he did for /dréait/, or, had he used *i* for /éi/, he would have written *drit* for /dréait/. The simplest and most reasonable explanation for graphs like *savir*, *cum*, etc. is that the scribe heard phonemically simple vowels /é/ and /ó/; if there was any incipient diphthongization, as suggested e.g. by Wallensköld 94-5, it could not have been phonemically significant or the scribe would have written *ei* and *ou*, *ie* and *uo*. Faced with the problem of symbolizing stressed tense /é/ and /ó/, he chose the letters *i* and *u*, perhaps because *e* and *o* were pronounced open, especially in church or other formal pronunciation of Latin.

The variation in the texts, in final unstressed position, between *e* and *a* (*fradre*, *fradra*), *e* and *o* (*Karle*, *Karlo*), and *o* and *u* (*deo*, *deus*), clearly points to an attempt on the part of the scribe to symbolize a vowel for which he had no orthographical tradition, presumably mid-central /ə/ (thus, for instance, Melander). To assume that the dialect of the Oaths preserved final /e o u/ intact (Tabachovitz *Etude*, ch. 3; Suchier, von Wartburg) requires one of two corollary assumptions. Either we must assume that every final *a*, *e*, *o*, or *u* must be interpreted literally as standing for /a/, /e/, /o/, or /u/ respectively, and hence that (say) final /e/ in P.Rom. /frátre/ could have given both */frádre/ and */frádra/ indifferently—an assumption contrary to the fundamental assumption of historical linguistics; or else we must (with Tabachovitz) make extensive emendations in the text, regularizing *fradra* to *fradre* etc.—though virtually all work done on the Oaths in the last half-century shows that we need to emend it less, not more, than is usually thought. In *meos* 'my' (nom. sg. m.) we probably have /méas/, as shown by the later development to ONFr. /mís/ (through */miéas/); probably likewise *deo* 'God' (obl.) stands for /déa/ and *deus* (nom.) for /déas/, though corresponding to these words ONFr. has the Latinisms *D(i)eu*, *D(i)eus*. By virtue of these considerations, we must interpret other instances of *e*, *o*, and *u* in final syllables, and probably *e* in other unstressed syllables also, as standing for /ə/, as in *poblo* /póblə/ 'people', *fazet* /fácət/ 'might do', *Karlus* /kárłəs/ 'Charles', *returnar* /rəturnár/ 'turn back'. Similarly for final -*a* which has arisen as a supporting vowel, e.g. *sendra* /séndrə/ 'lord', parallel in spelling to *fradra* /frádrə/

'brother'. But final *-a* coming from VL final /a/ and appearing in OSFr. as *-a* /a/ and in ONFr. as *-e* /ə/ is clearly to be read as /a/, e.g. *cosa* /kósa/ 'thing'.

2.2. CONSONANTS: /p b f v m n l r/, spelled with the same letters as in our transcription, and each with the single allophone represented thereby in the IPA; and /t d k g c ġ s/ with the following spellings and allophones:

/t/ *d* in word-final position in *plaid*, *t* elsewhere: [t]: /pláit/ *plaid* ['plaɪt] 'agreement'; /retornár/ *returnar* [rətor'nar] 'turn back'; /konsérvat/ *conservat* [kon'servat] 'he keeps'.

/d/ *dh* or *d* between vowels, *d* elsewhere; [ð] between vowels and perhaps (in view of ONFr. /fræðrə/ *fredre* 'brother' etc.) between vowel and [r] followed by vowel, [d] elsewhere: /kadúna/ *cadhuna* [ka'ðuna] 'every'; /podéar/ *podir* [po'ðer] 'to be able'; /dóanar/ *dunat* ['donat] 'gives'; /prendrái/ *prindrai* [pren'drai] 'I shall take'.

/k/ *K* in *Karl-*, *qu* before *e* and *i* and before *a* in *quant* and *nunqua(m)*, *ch* in *christian*, and *c* elsewhere: [k]: /kárle/ *Karle, Karlo* ['karlə] 'Charles'; /nónka/ *nunqua(m)* ['noŋka] 'never'.

/g/ *g*: perhaps [ɣ] between vowels and between vowel and [r] followed by vowel (by parallelism with /d/ and in view of ONFr. /saíremént/ *sairement* 'oath' < [saɣra'ment]): /sagráment/ *sagrament* [saɣra'ment] 'oath'. Presumably [g] elsewhere, though not attested in the Oaths.

/c/ *c*, *z*: [tʃ]: /fácet/ *fazet* ['fatʃet] 'that he may do'; /céast/ *cist* ['tʃest] 'this' (obl. m. sg.).

/ġ/ *j*: [ʒ]: /ġurát/ *jurat* [ʒu'rat] 'he swore'.

/s/ *s*: perhaps [z] between vowels (parallel to the voicing of VL unvoiced intervocalic plosives as in /sagráment/ < /sakramentu/, /póblə/ < /pópulu/), and [s] elsewhere: /kósa/ *cosa* ['kɔza] 'thing'; /séat/ *sit* ['set] 'it may be'; /nəúls/ [nə'uls] 'no one'.

On the basis of etymological evidence (VL /sénior/ > *sendra*) one is tempted to postulate also a phoneme /ń/, but there is no evidence for it in the orthography, and *sendra* may just as well stand for /séndrə/ as for /séndrə/.

2.3. STRESS must be postulated as phonemically significant (symbolized by /'/) on the basis of its having been significant both in VL and in later OFr., although it is not indicated in the orthographical tradition of the Oaths. There must have been a minimal contrast in verb forms ending in *-at*, e.g. /ġurát/ 'he swears' (< VL /júrat/ and > ONFr. /ġúrəθ/ *juret*) and /ġurát/ 'he swore' (< VL /júrát/ and > ONFr. /ġuráθ/ *jurat*), which would both have been spelled *jurat* in the Oaths; in this instance, it is only by the evidence of the OHG equivalent *gesuor* that we know the form is to be interpreted as /ġurát/, past tense.

2.4. JUNCTURE. We can postulate at least two types of juncture because of special treatment of sounds in the spelling:

1. Word-final disjuncture, with unvoicing of word-final voiced consonants, is present in /éant/ *int* 'from it' (< VL /éande/ < PRom. /inde/), and is indicated by the reverse spelling *d* for what could only have been /t/ in /pláit/ *plaid* 'agreement' (< VL /plákitu/ and > ONFr. /pláit/).

2. Close juncture between words in a phonemic phrase is indicated by such spellings as *dist* /déast/ 'from this' = *d* + *ist*; *lint* /léant/ 'him from it' = *l(o)* +

int. Certain consonant assimilations had clearly taken place in close juncture, e.g. /déft/ *dift* < */dévt/ < /débet/; perhaps, likewise, *Lodhuigs* is to be interpreted as /loduvíks/.

3. TENTATIVE PHONEMIC TRANSCRIPTION. We assume slow pronunciation (as would have been likely in a public assembly such as the one in which the Oaths were pronounced), and hence fully stressed individual words except in the case of prepositions, conjunctions, and some negatives and possessives; we do not indicate close juncture between the elements of what was probably a phonemic phrase except where such juncture is shown in the orthographical tradition. Since nothing can be inferred as to intonation patterns, we omit all punctuation except a period at the end of each Oath.

3.1. OATH OF LOUIS THE GERMAN: pro déa amóar e pro kristían póbla e nóstra komún salvamént déast dí en avánt en kánt déas savéar e podéar mæ dóanar sí salvarái éa céast mæon frádræ kárlæ ed en aiúda ed en kadúna kósa sí kóam óm pær dréait son frádræ salvár déaft en ó kéad íl mí [méa?] altresí fácæt ed a lodér núl pláit nóanka prendrái kí mæon vól céast mæon frádræ kárlæ en dálnæ séat.

3.2. OATH OF CHARLES THE BALD'S SOLDIERS: sí loduvíks sagramént kæ son frádræ kárlæ gúrat konsérvat e kárlæs méas séndre dæ sóa párt non læ s tánæt sí íó ræturnár non léant póis næ íó næ næúls kui éa ræturnár éant póis en núla aiúda kóntra loduvíg nóan li iv ér.

4. CLASSIFICATION OF THE LANGUAGE OF THE OATHS. Most discussion of the Oaths of Strassburg has proceeded on the assumption that their dialect was already Old French, and attempts at localization have been made on the basis of the identification of one or another characteristic of the Oaths with later dialectal features. Our discussion and phonemic interpretation would tend to show, however, that such attempts are nearly futile. The Oaths do not show any of the specific characteristics of Old North French, such as vowel-raising and diphthongization in PGRom. free syllables, or reduction of unstressed /a/ to /ə/, or palatalization of /k/ before /a/, etc. They do not show such specifically Old South French (Provençal) features as diphthongization or raising of vowels in the vicinity of palatals, nor the SEFr. preservation of final /e/ and /o/ as 'supporting vowels' which Meyer-Lübke believed they manifested. The only local feature that they seem to show is the Lorraine form *tanit* < /ténjat/ (Tabachovitz *Etude*). They do show the 'West Romance' sound shift, which they have in common with both ONFr. and OSFr., but not with PGRom. (cf. Hall 14-5), and which consequently marks the Oaths as belonging to a later stage than PGRom. Our interpretation of *i* and *u* (< VL /éa/ and /óa/) as standing for undiphthongized vowels removes most of the supposed Latinisms which some critics have seen in the Oaths, leaving only the *mm* of *commun* and the final *m* of *nunquam*. Between PGRom. and the ONFr. and OSFr. literary languages, we must assume an intermediate stage showing the West Romance sound shift which is common to both of these, but showing none of the features peculiar to them individually (as implied though not stated clearly enough in Hall 9-15), labelling it simply Pre-French. The dialect of the Oaths of Strassburg is extremely close to this stage, and we might characterize them as having been written in nearly undifferentiated, conservative Pre-French.

SEMANTIC DISTRIBUTION IN GAELIC DIALECTS

MYLES DILLON

Dublin Institute for Advanced Studies

It is a matter of frequent occurrence that words etymologically identical appear in varying forms in a given language. The pairs *person* : *parson*, *royal* : *regal*, *shirt* : *skirt* are familiar examples in English. In the first and third there is a clear difference of meaning, while in the second the distinction is rather one of style. The source of these doublets is different in each case. In the first a sound-change which has resulted in the modern standard English pronunciation of *clerk*, *Berkeley*, *Hertford* affected likewise *person*. But the change was not completed. The changed form, now written *parson*, has been retained only for the incumbent of a parish, while the common form *person* serves the other meanings.¹ In the second we have an instance of learned borrowing, while the third presents Scandinavian and English forms side by side, with a notable differentiation of meaning. The first of these three types is of peculiar interest, for the language, in such cases, has added to its vocabulary from within its own resources, not by means of suffix or composition, but out of the working of phonology. It is as though the speakers saw an opportunity, and availed themselves of it.

In French the case of *chaise/chaire* is well known, and has been explained by the reaction of educated people against a pronunciation regarded as vulgar, the superior form being then reserved for the pulpit. In areas where the change *chaire* > *chaise* was not regular, where *chaire* thus remained as the name for a common chair, the form *chaise*, associated with visitors from Paris, came to be used for the pulpit.²

My friend J. Orr has pointed out to me that the French pair *plier/ployer* provides another example of at least partially distributed meaning. Here the doublets arise, however, merely by double conjugation due to analogy. The treatment of final *-eur* and *-eux*, which fell together, has given rise to some further examples; thus, beside *faucheur* 'reaper', a 'back-formation' *faucheux* arose, and this survives as the name of the field spider.³

I am not aware that a linguistic term has been suggested for this process of semantic distribution. A number of examples in Irish and Scottish Gaelic have come to my notice, and my purpose here is to present them for consideration, in the hope that they may be of interest to linguists and lead to the observation of the process in other languages.

The treatment of Irish *ē* before non-palatal consonants in Scottish Gaelic is not uniform. O'Rahilly observes that Scottish Gaelic has a number of words in which an historic *ē* followed by a non-palatal consonant has acquired the value of *ao*;⁴ examples are *aodann* 'face' (Ir. *éadan*), *aodach* 'cloth' (Ir. *éadach*), *faodail*

¹ See H. C. Wyld, *History of modern colloquial English* 212 ff. (Oxford, 1936).

² W. von Wartburg, *Problèmes et méthodes de la linguistique* 24 (Paris, 1946).

³ *BSL* 48.8 (1951).

⁴ T. F. O'Rahilly, *Irish dialects past and present* 32 (Dublin, 1932).

'waif, thing found by chance' (Ir. *éadail*), *taod* 'halter' (Ir. *téad*). The conditions under which this development took place have not been established. Normally Ir. *ē* in this position remains in Scottish Gaelic or suffers diphthongisation to *ia*.⁵ There are some words in which the varying treatment of this sound has resulted in semantic distribution. Speakers from Lewis, North Uist, and Skye tell me that the form *taod* [tλ:ɔ̃] is reserved for 'halter', while *teud* [t'e:ɔ̃] means 'harp string';⁶ *faodail* 'waif' : *eudail* 'treasure, darling'; *sao*d [sλ:ɔ̃] (in *Dé'n saod tha ort?* 'How are you?') : *seud* [ʃe:ɔ̃] 'way' (Ir. *séad*). On the other hand, *deur* [d'e:r] means 'tear', while the diphthongised form *diar* has the wider meaning 'drop'.⁷ There is question of a similar distribution of meanings for *eud/iad* 'zeal, eagerness' / 'jealousy', but the evidence available to me is not clear.

We find also an alternation *ō/ua* with distributed meanings in Scottish Gaelic: *bòidheach* 'beautiful' : *buadhach* 'gifted, courteous'; *lòghmhar* 'excellent' : *luachmhar* 'precious, valuable'. Here in each case the second form seems to be literary, with restricted meaning. For the semantics one might compare *royal/regal*, although the linguistic history is quite different. The word *fàille* 'joy', 'welcome' seems to have had a by-form *faille* from a very early time, and both forms occur in Irish in both senses;⁸ but in Scottish dialects *fauille* (< *faille*) means 'good cheer', while *fàille* means 'welcome'.⁹

In Middle Irish *ō* was apparently an open vowel tending towards *ā*, and the Grammatical Tracts¹⁰ permit varying forms for a number of words in modern bardic poetry: *córuid cáraid*, *óinmhid áinmhid*, *tóithim táithim* II §13; *dóigh dáigh*, *fóir fáir* 14; *tinól tindl* 35; *friothólamh friothálamh* 37, 100, III 70; (*friothóilimh friotháilimh* II 150; *friothóileamh friotháileamh* II 101, III 70;) *anóir andáir*, *cóir cáir*, *égoir égcáir* II 42; *bóthar báthar* 53; *ród rád*, *fód fád* 88; *glór glár* 96; *iargnó iargná* 108; *léghóid légháid* 112; *térnódh térnádh* III 11; *cóirged cáirgeadh* 54;¹¹ *dígbódh dígbádh*, *fadódh fadádh*, *éltódh éltádh*, *élódh éladh*, *fasdódh fasdádh*, *claechlódh claechládh*, *síthlódh síthládh*, *tinntódh tinntádh*, *inntódh inntádh*, *impódh impádh*, *cadódh cadádh*, *cumnódh cumnádh* 64;¹² *in chóra in chára* IV 29.¹²

This list includes *cóir cáir* 'just, proper, correct' and *fód fád* 'sod'. Matheson tells me that in Lewis *ba chóire* 'it would be more generous' is distinguished from *ba chàra dhuit* 'it would be more fitting for you, you should'; *fàd* 'sod of turf' is distinguished from *fòd* in *fò'n fhòid* 'in the grave'. Here again the second is clearly the cultural form, just as Fr. *chaire* (or *chaise*) is cultural.

The Middle Irish *cerd* 'craft' appears in South Uist in two forms: *ceard* 'tinker'

⁵ Ib. 194, where O'Rahilly says that the change to *ia* is more frequent in the northern dialects.

⁶ *Teud* is therefore a learned form.

⁷ The Irish *deór* serves both meanings.

⁸ See *R.I.A. dictionary* F 24 (Dublin, 1950); O'Rahilly, *Irish dialects* 36.

⁹ For many of these examples, and some which follow, I am indebted to my friend Angus Matheson of the University of Glasgow.

¹⁰ *Eriu* 8-10, 14 (Dublin, 1916-46).

¹¹ These verbal nouns are also allowed forms in *-ádh* and should be kept apart.

¹² It may be observed that the conditions under which the variant form with *á* arises are not clear. Most words do not show it; e.g. *annró*, *bó*, *bróg*, *brón*, *gó*, *ól*, *trócaire*. But beside *dóibh* 'to them', *dáibh* is common in verse.

and *ciùird* (originally acc. and dat. sg.) 'trade'. Here an inflected form has supplied the second word. A similar case is that of Irish *sull* 'merriment', which appears in Lewis as *sull* 'obesity' and *solla* 'placid': *duine sultmhar* 'a heavy (fat) man'; *duine solta* 'a placid man'.¹³ The form with *o* is simply the genitive with differentiated meaning.

A more curious example is that of *soitheach* 'vessel', which in Modern Irish may be used of a sailing vessel or of a vessel to contain liquid. This is also the usage in Lewis and Harris. Dinneen records a variant *saothach* for parts of Munster, and in South Uist and Barra both forms occur with distributed meanings: *saothach* 'sailing vessel', *soitheach* 'drinking vessel'.¹⁴ Now in Lewis the Old Irish *ethar*¹⁵ has survived as *eathar* in the sense 'boat', but informants from Elgol in Skye were reported to me as using *eathar* for both 'pail' and 'boat'.¹⁶ It would seem, therefore, that in Skye *eathar* has borrowed the second meaning of *soitheach*, a process of semantic change for which analogies are to be found in Spanish. Castro has shown that some Spanish words have acquired the semantic range of Arabic words to which they were at first only partially equivalent.¹⁷

One example common to Scotland and Ireland is *aire/faire*. Here *aire* 'heed' has developed a prosthetic *f*, as commonly (e.g. OIr. *anaid* 'waits': Mod. Ir. *fanaim*; MIr. *uindeóc* 'window': Mod. Ir. *fuinneóg*, Sc. Gael. *uinneag*¹⁸). But the new form *faire* is reserved for special meanings: *aire dhuit!* 'be careful!'; *tabhair aire dhom* 'pay attention to what I say': *bhíos ag faire air* 'I was watching him'; Sc. Gael. *ag gabhail faire* 'keeping watch'.

More interesting is the opposition of *tugaim* 'I bring, give' to *tuigim* 'I under-

¹³ Angus McIntosh has told me of the analogous semantic development of English *buxom* (cf. G *biegsam*).

¹⁴ See C. H. Borgström, *Dialects of the Outer Hebrides* 141 (Oslo, 1940).

¹⁵ The word is discussed by O'Rahilly, *Celtica* 1.160 (1946). His etymology is, of course, not affected by my observations.

¹⁶ Farquhar Macintosh, a former student of mine in Edinburgh, has kindly sent me the following note:

Eathar was used to mean 'a pail, a bowl, a basin' in my own village of Elgol by two [sisters] now dead, Margaret MacKinnon, died c. 1950, age 82 years, and Kate MacKinnon, died 1952, age 88 years. ...

Mr. Lachlan Robertson, a colleague of mine in Glasgow Academy, who also comes from Elgol, confirms the use of *eathar* in the sense I have indicated. He also heard it from the two sisters I have mentioned. My father's cousin, John F. Macintosh, who is now about 70 years of age and who is also a native of Elgol, tells me that he often heard his mother use *eathar* in the sense of 'container', more specifically to mean 'an ewer'. ...

Angus Matheson has discovered an interesting confirmation of this use in the following stanza from a late-18th-century MS in the library of Glasgow University (MacLagan MS No. 159): 'N oiche bha mi'n Coire-mheangain / Bu bhochd m'ainis 's mo dhoidh; / B'i mo chluasag an sneachda / 'S b'e mo bhreacan a ghaoth reot, / 'S nan tarladh orm paethe / B'e m'eithir mo bhrog' 'The night that I was in Coire-mheangain, wretched was my plight: the snow was my pillow, and the freezing wind my plaid; and if thirst should come upon me, my shoe was my drinking-vessel.' (In MacLagan MS No. 166 the last line reads 'S e m'eathur mo bhrog; in MacLagan MS No. 70, B'e mo saochach mo bhrog.)

¹⁷ See A. Castro, *España en su historia* (Buenos Aires, 1948), reported in *Nueva revista de filología hispánica* 3.298-300. See also the discussion by Spitzer, *ibid.* 141, and Castro's reply, 149.

¹⁸ See C. Sarauw, *Irsk studier* 16-7 (Copenhagen, 1900); H. Pedersen, *VKG* 1.435 (Göttingen, 1909).

stand', also common to Scottish and Irish dialects. In Old Irish *do-ucc(a)i*, *do-uicci* serve both meanings ('can bring', 'understands'),¹⁹ the palatal *-cc-* having spread from forms like pret. 3 pl. *do-uicset* where *i* has been syncopated.²⁰ But in the later language, just as in French *plier/ployer*, two paradigms develop, with clearly distributed meanings.

Doublets with distributed meanings may rise by re-formation of compounds. M. A. O'Brien tells me that in Co. Donegal *sean-duine* when pronounced [ʃaNyn'ə] means 'old man' in opposition to *sean-bhean* 'old woman'; but the re-formation pronounced [ʃaNdyn'ə], without assimilation, has the literal meaning 'old person'.²¹

Finally there is an interesting example which appears only in Irish dialects, and which O'Brien has suggested to me. In Old Irish the dual personal numeral is *dias*, which remains in Scottish Gaelic as *dithis*, and to some extent in Co. Donegal as *dís*. In southern Irish dialects, however, *dias* is now obsolete, and the form *beirt* is used (*beirt fhear* 'two men', *beirt bhan* 'two women'). The word is best explained as originally dat.-acc. sg. of *beart* (OIr. *bert*) 'bundle'. It appears in the literature rarely and only late, although it has largely ousted *dias* from the living dialects. My only example from a text (supplied by Brian Ó Cúfá) is *beirt ridireadh* 'two knights', *Cín Lae Ó Mealláin* 37.6 (c. 1646 A.D.).²²

Giacomo Devoto has pointed out to me that these examples of semantic distribution show sound laws operating not mechanically but teleologically. Gilliéron contended long ago that sound laws did not work, and must be abandoned. The majority of linguists still cling to them, and I should hate to have to operate without them. But it is no treason to admit that, in spite of the sound laws, each word has its own history. Some words, fertilized by the sound laws, throw off shoots which may strike a fresh root and flourish independently.

¹⁹ See Pedersen, *VKG* 2.471-2 (Göttingen, 1913).

²⁰ See Thurneysen, *Grammar of old Irish* 103 (Dublin, 1946).

²¹ In Donegal, beside *deichneabhar* 'decade' (of the rosary), there is a new form *deichear* 'a group of ten' by analogy with *seachtar*, *ochtar*. In Waterford, beside *deachmhadh* 'tithe', there is a new analogical form *deicheadh* 'tenth' (Sheehan, *Sean-Chaint na n Déise* 123 §18 [Dublin, 1944]).

²² *Analecta hibernica*, No. 3 (Dublin, 1931).

RUSSIAN DECLENSIONAL MORPHEMES

GEORGE L. TRAGER

Foreign Service Institute

0. The purpose of this article is twofold: to examine the various patterns of declension found in Russian and establish the MORPHEMES in them; and, by this presentation, to exemplify a possible method of morphemic analysis.

The procedure involves the identification of the MORPHS that occur, and their assignment as ALLOMORPHS to morphemes. The methods used are summarized and exemplified in the author's treatment of English.¹ Some additional explanation is in order. The forms that occur are examined in phonemically isolated words (§0.1). They are compared and then segmented mechanically at points that leave partials which recur in more than one form. Thus the forms ruká (see §0.3 for the orthography), rúku, rúki, rukáx enable us to segment into ruk-, ú , ú , and the partials -a, -u, -i, -ax; these can be compared with nogá, nógu, nógi, nogáx, and are thus established as recurring; further comparison of rukáx, rukám, rukámji establishes the recurring partial -a-, and the partials -x, -m, -mji; and the latter are established as recurring by comparison with nóvix, nóvim, nóvimji. After a tentative set of morphs has been established in this way, the forms are tested in comparable frames and sequences to determine which have identical functions: thus odná + ruká, odná + kóstj compared with odnòy + rukí, odnòy + kóstji, establish the functional identity of -a in odná and in ruká, and of these instances of -a with the absence of a final vowel in kóstj, this absence being set up as a zero element, -0.^{1a} We can then say that there is a morpheme which contains at least the allomorphs -a and -0; the morpheme can be symbolized by some arbitrary symbol, say $\sqrt{-N}$ (see §0.3). That our segmentation is tentative can be shown by the fact that from odnòy + rukí and odnòy + kóstji alone we could only conclude that -oy and -i are allomorphs of a single morpheme, whereas the introduction of further evidence makes it clear (as will be shown below) that odnòy has to be analyzed as odn-0-óy-0, and that it is the final -0 that goes with -i in a morpheme $\sqrt{-G}$. In this connection it is emphasized that no zero morphemes exist, but only zero allomorphs of morphemes whose other allomorphs are not zero. Stress patterns are separate morphs, forming suprasegmental morphemes.

The constant aim in this procedure is to obtain the smallest possible number of different morphemes, with one morpheme for each category throughout the analysis, and to arrive at the results wholly by inspection and commutation in similar structural frames, without the use of meaning. Since the structural

¹ George L. Trager and Henry Lee Smith Jr., *Outline of English structure* 53 ff., esp. §§2-2.13 (SIL Occasional papers, No. 3, 1950).

^{1a} For 'zero' the author prefers a more distinctive symbol, like that used in the work referred to above (fn. 1): a zero with a diagonal crossbar slanting downward from left to right. This is replaced here, for typographical reasons, by an ordinary zero. For similar reasons, the symbol introduced below (§0.3) for 'allomorph'—a root sign with exponent q —is here used in place of the author's original choice, a root sign with exponent a .

frames used are not themselves analyzed but merely determined to be valid, the analysis and the segmentation remain wholly in the realm of morphology, the division of morphemics (as I use this term) dealing with the structure of words.²

0.1. Russian morphology involves the following entities. A word is defined as consisting of a BASE and a SUPERFIX; it may include one or more SUFFIXES. The suffixes may be STEM-FORMING (derivational) or INFLECTIONAL; inflectional suffixes are PREFINAL and FINAL (rarely PRE-PREFINAL). Superfixes are principally of the types $\sqrt{\text{...}}$ (initial primary stress, with one to four weak syllables following; $\sqrt{\text{...}}$ is to be read 'the morpheme') and $\sqrt{\text{...}^{\text{...}}}$ (final primary stress), including specific forms such as $\sqrt{\text{...}}$, $\sqrt{\text{...}}$, and the like.

Russian words that have inflectional suffixes occur in sets that constitute PARADIGMS; each constituent of a paradigm is a word, but as a member of such a set it is called a PARALOG. The morphemes that constitute a set of inflectional suffixes in a paradigm are called PARAMORPHEMES. A paralog, then, is a base combined with a paramorpheme. According to the paradigms into which they enter, Russian words may be grouped in the so-called parts of speech. The parts of speech to be considered here are similar in their inflection (called declension) and constitute the over-all class of ENTITIVES, comprising NOUNS, ADJECTIVES, PRONOUNS, and NUMERALS.

0.2. The categories of inflection are determined by the grouping of allomorphs in complementary distribution into morphemes. The several classes of entitives are defined by the morpheme categories which they include in their paradigms. Nouns are inflected for the cases (NOMINATIVE, ACCUSATIVE, GENITIVE, LOCATIVE, DATIVE, INSTRUMENTAL, PARTITIVE, LOCAL, VOCATIVE), and for PLURAL number. Adjectives are inflected for the first six cases and for plural number, and also for gender (MASCULINE, NEUTER, FEMININE). Numerals are inflected for the first six cases, with two instances of (feminine) gender inflection. Pronouns ('personal' and 'interrogative-relative' pronouns only) are inflected for the first six cases, with some instances of plural and one of gender. Nouns are classified as to gender (by syntactic commutation), but not inflected for it; nouns and adjectives are classified as animate or inanimate, these being characteristics of specific noun-stems and syntactic effects on adjectives. Numerals and pronouns are small limited classes, each member of which is identified by its specific metalinguistic reference (meaning).

Each of the categories of inflection is considered to be represented by a single morpheme. For each morpheme we must give a statement of the occurrence of its allomorphs in terms of the stems or classes of stems which it follows and of the other inflectional morphemes that are present. Stem-formation as such is not treated in this paper.

Noun-stems are listed as belonging to different types, depending on the paradigms they exhibit. The principal types are the FIRST, SECOND, and THIRD DECLENSIONS. Each declension has its own set of paramorphemes. There are also INDECLINABLE nouns, which belong to the class by virtue of being classified as to gender and having one case-morpheme, the nominative (singular). Some nouns are not classifiable as to gender, but they then exhibit the plural mor-

² Op.cit. 54-5.

pHEME. (The genders and the plural are mutually exclusive throughout the system, except possibly in one paradigm; see §2.4.)

Adjective-stems are similarly listed in groups, depending on the allomorphs of the ADJECTIVE-INFLECTION morpheme that occur (see §2.5; this morpheme has nothing to do with any derivational suffixes that may be present). The groups are: REGULAR adjectives, *y*-adjectives, *ov*-adjectives, and PRONOMINAL adjectives. Regular adjectives (but not all of these) are the only ones that have the so-called short forms; these forms are analyzed as having only gender suffixes, without an adjective-inflection morpheme (§§2.2-5).

Numerals and pronouns are specifically identified, as stated, by their bases.

It should be noted that 'singular' is not an inflectional category. Forms of entitives which are not plural—that is, do not have any allomorph of the plural morpheme—may be referred to as SINGULAR, which is then a category only in the sense of a basic form (an 'unmarked' form) from which others differ by having an explicit characteristic.

0.3. The several classes and categories delimited in §0.2 are referred to by the following abbreviations and morpheme-symbols:

I, II, III: first, second, third declension	'40': the numeral 'forty'
Sg, Pl: singular (basic or 'unmarked' category—no morpheme), plural	'100': the numeral 'hundred'
n., adj., pers.pr.: noun(s), adjective(s), personal pronoun(s)	✓-N: the nominative morpheme
M, Nt, F: masculine, neuter, feminine (also as morphemes, ✓-M-, ✓-Nt-, ✓-F-)	✓-A: the accusative morpheme
1, 2, 3, R: first, second, third person, reflexive	✓-G: the genitive morpheme
reg.adj.: regular adjectives (all except the following three groups)	✓-L: the locative morpheme
y-adj.: adjectives with the adjective-inflection allomorph ✓/y-	✓-D: the dative morpheme
ov-adj.: adjectives with the adjective-inflection allomorphs ✓/ov- and ✓/in-	✓-In: the instrumental morpheme
pron.adj.: pronominal adjectives—the words for 'one', 'this', 'that', 'my', 'thy', 'one's own', 'our', 'your', 'whose', 'self', 'all'	✓-Pt: the partitive morpheme
'2', '3', '4', 'both': the numerals 'two', 'three', 'four', 'both'	✓-L ₂ : the local morpheme
num. '5+': the numerals from '5' up, except 'forty' and 'hundred'	✓-Vo: the vocative morpheme
	✓-Pl: the plural morpheme
	✓-Adj: the adjectival-declension morpheme
	0: 'zero' allomorph
	A-, -A: the singular and plural elements of a stress paradigm in which the superfix is ✓ ² ...
	B-, -B: superfix ✓... ²
	C-, -C: superfix ✓ ² ... in SgA and PlN, ✓... ² elsewhere in the paradigm

Specific words are either cited (in a morphophonemic orthography) or identified by a gloss only, as 'thousand'.

All morphemes are identified by the symbol ✓, and allomorphs by ✓/. Bases are shown with a terminal hyphen, as ✓s(o)t- (the symbol in parentheses is the 'movable vowel'); inflectional prefinals have initial and terminal hyphens, as ✓-Pl-; finals have an initial hyphen, as ✓-N; word superfixes have a hyphen under the primary stress mark. Morphemes and morphemically analyzed words are written with phonemic symbols (see next paragraph) used morphophonemically, as ✓on- '3 pers.pr.', phonemically /on-/ or /an-/ depending on the superfix.

The principal morphophonemic writings are the following: the vowels \sqrt{o} and \sqrt{e} are set up in morphemes that contain the phonemes /o/ and /e/ under the primary stress of a superfix in one or more forms of a paradigm, in alternation with /a/ (or /i/) and /i/ respectively under weak stress in other forms; voiced final or ante-voiceless consonants (for voiced-voiceless pairs) are set up when voiced consonant phonemes occur in nonfinal or ante-vocalic position in other forms of a paradigm.

The phonemic symbols used are /p t k b d g c č j f s š x v z ž γ m n l r y/, and /j/ as the symbol for palatalization (without arguing whether this is a phoneme or a component). Not all the phonemes occur in the examples. Stresses are primary /'/, weak /˘/ (often unmarked), and medial /˘/ (rare in the examples). Internal juncture is marked /+/ (plus). Terminal junctures and pitch phonemes do not enter into this discussion.

1. The five case morphemes $\sqrt{-N}$, $\sqrt{-G}$, $\sqrt{-L}$, $\sqrt{-D}$, $\sqrt{-In}$ are found to occur, as discrete forms, in the Sg and with $\sqrt{-Pl}$, in all nouns of declension II. In declension I the L and D are always alike in the Sg; in III the G, L, and D are alike; indeclinable nouns have all forms alike, in both Sg and Pl; all adjectives, and the third-person pronoun, have the G, L, and D alike in the SgF, and the G and L alike in the Pl; the pronouns 1 and 2 in the Pl have G and L alike; the numerals '2', '3', '4', and 'both' have G and L alike; the numerals '5+' have G, L, and D alike; '40' and '100' in the Sg have G, L, D, and In alike. The statements just made are cast in traditional form; they might be restated to indicate that there is no difference, in the paradigms mentioned, between the forms that occur where the nouns of II have the case forms named. This could then be interpreted to mean that one or more of the cases simply do not occur in the paradigms named—for instance, that D does not occur in the Sg of I; this is the treatment we adopt for the cases other than these five (below), and for all other instances of homonymy within different syntactic frames. However, because of the statistical preponderance of the nouns of II, we have elected to regard the five cases occurring there as occurring in all declined forms; but indeclinable nouns and '40' and '100' [Sg] are not included in the class for this purpose. Thus we set up some allomorphs as homonymous with allomorphs belonging to other morphemes.

The $\sqrt{-A}$ occurs only in declension I, and in association with the F gender in adjectives, one pronoun, and two numerals (§2.4). The $\sqrt{-Pt}$ occurs only with some nouns of declension II, which have to be listed; similarly, $\sqrt{-L_2}$ occurs only with some nouns of II and III. The $\sqrt{-Vo}$ occurs only with five nouns ('Father' [i.e. God, not the common noun 'father'], 'God', 'Lord', 'Christ', 'Jesus'). The case morphemes are mutually exclusive; each occurs last in a word, after all other inflectional morphemes.

In the forms that may be called Sg, the selection of case allomorphs is conditioned by the declension to which a noun-stem belongs or (in adjectives) by the gender morphemes. Where $\sqrt{-Pl}$ (§2.1) is present, the case allomorphs are conditioned by it.

Nouns classified as animate have the $\sqrt{-G}$ with the $\sqrt{-Pl}$ where the $\sqrt{-A}$ is syntactically expected, while inanimate nouns have the $\sqrt{-N}$ in that situation.

In II, animates have $\sqrt{-G}$ for expected $\sqrt{-A}$ in the Sg also, but in III they have $\sqrt{-N}$. All inanimates have $\sqrt{-N}$ for $\sqrt{-A}$ in the Sg. Adjectives, and numerals in certain syntactic sequences, have $\sqrt{-G}$ for $\sqrt{-A}$ where syntactically conditioned as animate by animate nouns. All nouns of I, and F adjectives, have the $\sqrt{-A}$ in the Sg. The pronouns have $\sqrt{-G}$ for expected $\sqrt{-A}$, except 'what', which has $\sqrt{-N}$.

1.1. The nominative morpheme ($\sqrt{-N}$) has four allomorphs: $\sqrt{-a}$, $\sqrt{-0}$, $\sqrt{-o}$, $\sqrt{-l}$ ('loss of stem-final consonant or weak syllable'). Their distribution is:

$\sqrt{-a}$: Sg I n.; Sg F adj. and 3 pers.pr. (Examples: ruk-á 'hand', kúxnj-a 'kitchen'; nóv-0-ay-a 'new', on-0-á 'she'.)

$\sqrt{-0}$: Sg II M (except as listed) n.; Sg III M, F (except as listed) n.; listed indeclinable n. (with stems ending in consonants or in vowels, of any gender); Sg M adj. and 3 pers.pr.; 1 and 2 pers.pr.; num. '5+'; '40'; with $\sqrt{-Pl}$ everywhere; with $\sqrt{-F}$ in '2', 'both'. (stól-0 'table', d(jé)nj-0 'day'; pútj-0 'path' [the only M in III], kóstj-0 'bone'; kingurú-0 M 'kangaroo', intjirvjyú-0 Nt 'interview', madám-0 F 'madam(e)'; nóv-0-iy-0 'new'; ón-0-0 'he'; yá-0 'I', tí-0 'thou'; pjátj-0 'five'; sórok-0 'forty'; rúk-i-0 'hands', stol-i-0 'tables', putj-i-0 'paths', kóstj-i-0 'bones', nóv-iy-a-0 'new [ones]', onj-i-0 'they', m-i-0 'we', v-i-0 'you'; dv-á-0, dv-jé-0 [F] 'two', ób-a-0, ób-je-0[F] 'both'.)

$\sqrt{-o}$: Sg II Nt (and listed M) n.; listed indeclinable Nt n.; Sg Nt adj. and 3 pers.pr.; 'who' and 'what'; '100'. (okn-ó 'window', mórj-o 'sea', gorodjíšč-o M 'large town'; paljt-ó 'coat'; nóv-0-oy-o, on-0-ó, kt-ó, št-ó; s(o)t-ó.)

$\sqrt{-l}$: Sg III Nt n.; 'mother', 'daughter' (both III F). (/ímja/ 'name' = ímjónj-l, /djítj-á/ 'child' = djítj-átj-l; /mátj/ 'mother' = mátjirj-l, /dóč/ 'daughter' = dóčirj-l.)

1.2. The accusative morpheme ($\sqrt{-A}$) has the allomorph $\sqrt{-u}$ everywhere except in one instance, where it has $\sqrt{-o}$:

$\sqrt{-u}$: Sg I n.; Sg F adj. (except 'self'). (rúk-u, kúxnj-u; nóv-0-uy-u.)

$\sqrt{-o}$: Sg F of 'self'. (sam-0-oy-ó; the forms sam-0-uy-ú [as if reg. adj.] and sam-0-0-ú [as for other pron. adj.] also occur.)

As already stated, no other A forms occur; all other entitives use the G or the N form instead.

1.3. The genitive morpheme ($\sqrt{-G}$) has nine allomorphs: $\sqrt{-i}$, $\sqrt{-a}$, $\sqrt{-vo}$, $\sqrt{-o}$, $\sqrt{-0}$, $\sqrt{-ov}$, $\sqrt{-ey}$, $\sqrt{-x}$, $\sqrt{-s}$. Their distribution is:

$\sqrt{-i}$: Sg I, III n.; num. '5+'; '100' (after '2'). (ruk-í, kúxnj-i; kóstj-i, ímjónj-i; pjátj-i; dvjéstj-i, stj- being a special form of the stem here.)

$\sqrt{-a}$: Sg II n.; Sg M *ov*-adj.; Sg 1, 2, R pers.pr.; '40', '100'. (stól-á; djádj-in-0-a 'of uncle's...'; mjínj-á 'of me', tjíb-j-á 'of thee', sjíb-j-á 'of oneself'; sorok-á 'of forty', s(o)t-á 'of 100', /trjístá/ '300' [in the numerals '40' and '100' the $\sqrt{-G}$ occurs throughout the Sg where $\sqrt{-L}$, $\sqrt{-D}$, or $\sqrt{-In}$ is expected].)

$\sqrt{-vo}$: Sg M reg., *y*-, pron. adj., and 3 pers.pr.; 'who', 'what'. (nóv-0-o-vo 'of the new', trjéstj-y-o-vo 'of the third', moy-0-o-vó 'of mine', y-o-vó 'his'; k-o-vó 'of whom', č-o-vó 'of what'.)

$\sqrt{-o}$: Sg F 3 pers.pr. (y-oy-ó 'her'.)

$\sqrt{-0}$: Sg F adj.; with $\sqrt{-Pl}$ in listed I, II, III Nt n. (nóv-0-oy-0, trjéstj-y-

oy-0, t(ot)-0-óy-0 = /tót/ 'of that one', djádj-in-oy-0; rúk-0-0 'of hands', gláz-0-0 'of eyes', imjón-0-0 'of names'.)

√-ov: with √-Pl- in listed II n. (dom-0-óv 'of houses'.)

√-ey: with √-Pl- in listed I, II, III F n. (tjótj-0-ey 'of aunts', nož-0-éy 'of knives', kostj-0-éy 'of bones'.)

√-x: with √-Pl- in adj. and 3 pers.pr.; '2', '3', '4'. (nóv-0-i-x, trjéty-y-i-x, moy-j-i-x, djádj-in-i-x, y-i-x 'their'; dv-ú-x 'of two', trj-ó-x 'of three', četirj-ó-x 'of four'.)

√-s: with √-Pl- in 1, 2 pers.pr. (n-á-s 'of us', v-á-s 'of you'.)

1.4. The locative morpheme (√-L) has six allomorphs: √-je, √-i, √-m, √-0, √-x, √-s. The last one has wholly and the preceding two have partly the same distribution as the homonymous √-G allomorphs. The distribution is:

√-je: Sg I, II n.; Sg M *ov*-adj. as listed (when as proper names), Sg 1, 2, R pers.pr. The /j/ is lost automatically after a stem-final palatalized, palatal, or non-palatalizable consonant (i.e. Cj, Ć = /č j š ž y/, or K = /k g c x γ/). (golov-jé '[in the] head', ruk-jé = /ruké/, kúxnj-je, duš-jé '[in the] soul', dóm-je '[in the] house'; /a+pjitróvj/ = o+pjotr-óv-je 'about Petrov'; mn-jé, tjibj-jé, sjibj-jé.)

√-i: Sg III n.; num. '5+'. (kóstj-i, ímjónj-i; pjatj-i.) The L of nouns in /...iya/, having historical *i* in the orthography (*Rossii*), is best considered a spelling of √-je; otherwise √-i must be said to appear also in I and II n. when the stem ends in /...iy/.

√-m: Sg M adj. (except *ov*-adj. as listed, which have √-je) and 3 pers.pr.; 'who', 'what'. (nóv-0-o-m, trjéty-y-o-m, djádj-in-o-m, moy-0-ó-m, nj-ó-m [only the alternate stem √nj- appears in the L of the M 3 pers.pr., because the stem √y- is not used after prepositions and the L occurs only after a preposition]; k-ó-m, č-ó-m.)

√-0: Sg F adj. and 3 pers.pr. (nóv-0-oy-0, nj-éy-0 [see under √-m for the stem √nj-].)

√-s: with √-Pl- in 1, 2 pers.pr. (n-á-s, v-á-s.)

√-x: with √-Pl- everywhere else. (ruk-á-x, dom-á-x, nóv-0-i-x, dv-ú-x.)

1.5. The dative morpheme (√-D) has six allomorphs: √-je, √-u, √-i, √-mu, √-0, √-m. The first, third, and fifth are homonymous with allomorphs of √-L, which have the same (√-i and √-0) or partly the same (√-je) distribution. The distribution is:

√-je: Sg I n.; Sg 1, 2, R pers.pr. Loss of /j/ takes place as for √-je of √-L. (golov-jé '[to the] head', ruk-jé, kúxnj-je, duš-jé; mn-jé, tjibj-jé, sjibj-jé.)

√-u: Sg II n.; Sg M *ov*-adj. (dóm-u; pjotr-óv-u, djádj-in-u.)

√-i: Sg III n.; num. '5+' (kóstj-i, ímjónj-i, pjatj-i.)

√-mu: Sg M adj. (except *ov*-adj.) and 3 pers.pr.; 'who', 'what'. (nóv-0-o-mu, trjéty-y-o-mu, t-0-o-mú '[to] that [one]', y-o-mú '[to] him'; k-o-mú, č-o-mú.)

√-0: Sg F adj. and 3 pers.pr. (nóv-0-oy-0, y-éy-0.)

√-m: with √-Pl- everywhere. (ruk-á-m, dom-á-m, nóv-0-i-m, dv-ú-m.)

1.6. The instrumental has six allomorphs: √-oy(u), √-om, √-y(u), √-m, √-mji, √-mja. The /u/ of √-oy(u) is usually not present in colloquial Russian,

except in a few fixed expressions; the same is true of \mathcal{V} -y(u) whenever there is a preceding allomorph ending in a vowel. The distribution is:

\mathcal{V} -oy(u): Sg I n. (except 'thousand'); Sg 1, 2, R pers.pr. (ruk-óy(u), kúxnj-oy(u); mn-óy(u), tob-óy(u), sob-óy(u).)

\mathcal{V} -om: Sg II n.; Sg III M Nt n. (dóm-om; putj-óm '[by] the road', im-jonjom.)

\mathcal{V} -y(u): Sg III F n.; Sg F adj. and 3 pers.pr.; 'thousand' (Sg I F); num. '5+'. (kóstj-yu; nóv-0-oy-y(u) = /nóvay(u)/, y-éy-y(u) = /yéy(u)/; tísjič-yu; pjatj-yú '[with] five'.)

\mathcal{V} -m: Sg M adj. and 3 pers.pr. (nov-0-i-m '[with the] new [one]', y-f-m '[with] him').

\mathcal{V} -mji: with $\sqrt{\text{Pl}}$ everywhere except '2', '3', '4'. (ruk-á-mji, kúxnj-a-mji, dom-á-mji, kostj-á-mji, ludj-0-mjí '[with] people'; nóv-0-i-mji, t-j-é-mji, y-f-mji.)

\mathcal{V} -mja: with $\sqrt{\text{Pl}}$ in '2', '3', '4'. (dv-u-mjá, trj-o-mjá, četirj-0-mjá.)

1.7. The partitive morpheme $\sqrt{\text{Pt}}$ has only one allomorph, \mathcal{V} -u. It is found only with listed Sg II M n.; its actual occurrence suggests that as a category it goes in theory with all Sg II M n., and that where it is lacking, metalinguistic or even non-linguistic considerations are involved. The syntax of Russian allows the alternate occurrence of $\sqrt{\text{G}}$ in many or even most of the instances for which $\sqrt{\text{Pt}}$ is listed, and outside of Sg II it is always the $\sqrt{\text{G}}$ that occurs in the constructions that may be called partitive. (Examples of $\sqrt{\text{Pt}}$: sáxar-u '(some) sugar', naród-u 'people', čáy-u '(some) tea'.)

1.8. The local morpheme $\sqrt{\text{L}_2}$ has two allomorphs: \mathcal{V} -u, \mathcal{V} -i. Local forms always have the superfix $\sqrt{\dots}$, regardless of the superfix paradigm of a noun. The allomorphs are distributed as follows:

\mathcal{V} -u: Sg II M listed n. (v+ljes-ú 'in the woods', na+bjerjeg-ú 'on the shore'.)

\mathcal{V} -i: Sg III F listed n. (na+dvjerj-í 'at the door'—cf. na+dvjérji 'on the door', with $\sqrt{\text{L}}$.)

The $\sqrt{\text{L}_2}$ is replaced even for the listed occurrences by $\sqrt{\text{L}}$ as an alternant, and nouns outside of the stated types have $\sqrt{\text{L}}$ in constructions of this kind.

1.9. The vocative morpheme $\sqrt{\text{Vo}}$ occurs, as stated, in only five nouns; all others use forms with $\sqrt{\text{N}}$ in the 'vocative' construction. It has one allomorph, \mathcal{V} -je; the symbol \mathcal{J} represents a special palatalization that results in the morphophonemic changes /c/ > /č/ and /ɣ/ (or /g/) > /ž/, but otherwise yields ordinary palatalized consonants. The vocative forms found are: ót(je)c-je = /ótči/ 'Father' < ot(jé)c-0; bóɣ-je = /bóži/ 'God' < bóɣ-0; góspod-je = /góspadjí/ 'Lord' < góspódj-0; yisús-je 'Jesus'; xrijst-jé 'Christ' < xrijstós-0.

2. The case morphemes discussed in §1 may be preceded by the morphemes of the plural and the genders—masculine, neuter, feminine. The gender morphemes exclude the plural one (for a possible exception see §2.4). The allomorphs of the number and gender morphemes are conditioned by the declension to which a noun-stem belongs, or by the adjective-declension morpheme (or its absence, in short adjectives—see §2.5).

The plural morpheme occurs in nouns, in adjectives, in the numerals 'two', 'both', 'three', 'four', and in the personal pronouns of the first, second, and third

persons. The gender morphemes occur in adjectives, in the numerals 'two' and 'both', and in the third-person pronoun.

The adjective-declension morpheme is found in all adjectives (except the reg.-adj. short forms).

2.1. The plural morpheme ($\sqrt{\text{-Pl}}$) has six allomorphs: $\sqrt{\text{-i}}$, $\sqrt{\text{-0}}$, $\sqrt{\text{-a}}$, $\sqrt{\text{-e}}$, $\sqrt{\text{-u}}$, $\sqrt{\text{-o}}$. Their distribution is:

$\sqrt{\text{-i}}$: with $\sqrt{\text{-N}}$ in I, listed II, III F n.; *y-*, *ov-*, pron.adj. (except 'that' and 'all') and 3 pers.pr.; '3' and '4'; 1 and 2 pers.pr. (*rúk-i-0* 'hands', *stol-í-0* 'tables', *yáblok-i-0* 'apples', *kóstj-i-0* 'bones'; *trjétj-y-i-0* '[the] third [ones]', *djádj-in-i-0* '[those] belonging to uncle', *ét-j-i-0* 'these', *onj-f-0* 'they'; *trj-í-0* 'three', *četirj-i-0* 'four'; *m-í-0* 'we', *v-í-0* 'you'.) — with $\sqrt{\text{-G}}$, $\sqrt{\text{-L}}$, $\sqrt{\text{-D}}$, $\sqrt{\text{-In}}$ in all adj. (except 'that' and 'all') and 3 pers.pr. (*nóv-0-i-x*, *nóv-0-i-m*, *nóv-0-i-mji* '[of, about; to, with] new [ones]'; *y-í-x*, *y-í-m*, *y-í-mji* 'their, [about, to, with] them'.) — alone (i.e. directly after the stem, and without a following case morpheme) in short forms of reg.adj. (*nóv-i*.)

Forms like */gráždanji/* 'citizens' are analyzed (like *onj-f-0*) as containing plural $\sqrt{\text{-i}}$, the preceding palatalization being assigned to a special variant of the stem found only in the Pl N; if this turns out to be uneconomical in devising statements about stems (with which, as already noted, we are not dealing here; see also §4), an additional plural allomorph $\sqrt{\text{-ji}}$ will have to be set up, with distribution limited to */anjí/* and to nouns with Pl N */...a(')nji/* corresponding to Sg */...a(')nji(')n/*. A factor to consider here also is the existence of past forms of verbs with */lji/* in the Pl; in an over-all picture of Russian structure, these must be considered along with such formations as the short forms of adjectives. As indicated, the latter have $\sqrt{\text{-i}}$ in the Pl; the verb forms have a Pl formant $\sqrt{\text{-ji}}$; if an allomorph of $\sqrt{\text{-Pl}}$ in the form $\sqrt{\text{-ji}}$ is set up for the instances discussed just now, the verb Pl formant of the same shape can be regarded as actually the same form; it occurs in verbs without a preceding adjective-declension morpheme or a following case morpheme, as does $\sqrt{\text{-i}}$ in short adjectives.

$\sqrt{\text{-0}}$: with $\sqrt{\text{-G}}$ (whose different allomorphs are determined as listed for the specific stems—§1.3) in all n. (*rúk-0-0* 'of hands', *tjótj-0-ey*, 'of aunts', *dom-0-óv* 'of houses', *gláz-0-0* 'of eyes', *nož-0-éy* 'of knives', *kostj-0-éy* 'of bones', *imjón-0-0* 'of names'.) — with $\sqrt{\text{-In}}$ in listed III F n. [often freely alternating with $\sqrt{\text{-a}}$], and in '4'. (*ljudj-0-mjí*, *četirj-0-mjá*.)

$\sqrt{\text{-a}}$: with $\sqrt{\text{-L}}$, $\sqrt{\text{-D}}$, $\sqrt{\text{-In}}$ of all n. (except listed III F for $\sqrt{\text{-In}}$); 1, 2 pers.pr.; '100'. (*ruk-á-x*, *ruk-á-m*, *ruk-á-mji*, *tjótj-ax*, *dom-á-m*, *glaz-á-mji*, *nož-á-x*, *kostj-á-m*, *imjon-á-mji*; *n-á-s* '[about]us', *n-á-m*, *n-á-mji*, *v-á-s* '[about]you', *v-á-m*, *v-á-mji*; *s(o)t-á-x* '[about]hundreds'. — with $\sqrt{\text{-N}}$ in listed II and in III Nt n.; reg.adj.; '2'. (*dom-á-0* 'houses', *imjon-á-0* 'names'; *nóv-iy-a-0* 'new [ones]'; *dv-á-0* 'two'.) — with $\sqrt{\text{-G}}$ in 1, 2 pers.pr. (*n-á-s* 'of us', *v-á-s* 'of you'.)

$\sqrt{\text{-e}}$: with all cases in the pron.adj. 'that', 'all'. (*t-j-é-0* 'those', *t-j-é-x*, *t-j-é-m*, *t-j-é-mji*; *v(je)sj-j-é-0* = */fsjé/* 'all', *v(je)sj-j-é-x*.)

$\sqrt{\text{-u}}$: with $\sqrt{\text{-G}}$, $\sqrt{\text{-L}}$, $\sqrt{\text{-D}}$, $\sqrt{\text{-In}}$ in '2'. (*dv-ú-x*, *dv-ú-m*, *dv-u-mjá*.)

$\sqrt{\text{-o}}$: with $\sqrt{\text{-G}}$, $\sqrt{\text{-L}}$, $\sqrt{\text{-D}}$, in '3', '4', and $\sqrt{\text{-In}}$ in '3'. (*trj-ó-x*, *trj-ó-m*, *trj-o-mjá*; *četirj-ó-x*, *četirj-ó-m*.)

2.2. The masculine gender morpheme $\sqrt{-M-}$ has five allomorphs: $\sqrt{-oy-}$ ($\sim \sqrt{-iy-}$), $\sqrt{-o-}$, $\sqrt{-i-}$, $\sqrt{-e-}$, $\sqrt{-0-}$. They occur as follows:

$\sqrt{-oy-}$: with $\sqrt{-N}$ in reg. adj. (boljš-0-óy-0 'big'.) This allomorph has the variant /iy/ when under weak stress in the literary language, though even many standard speakers use /ay/ (automatically from $\sqrt{-oy-}$).

$\sqrt{-o-}$: with $\sqrt{-G}$, $\sqrt{-L}$, $\sqrt{-D}$ in reg., *y-*, and pron.adj.; 3 pers.pr.; 'who', 'what'; and with $\sqrt{-L}$ in non-listed *ov-*adj. (nóv-0-o-vo 'of the new [one]', nóv-0-o-m, nóv-0-o-mu, trjéty-y-o-vo, od(ji)n-0-o-vó 'of one'; y-o-vó, 'his, of him', nj-ó-m '[about] him', y-o-mú 'to him'; k-o-vó, 'whose', k-ó-m, k-o-mú, č-o-vó 'of what', č-ó-m, č-o-mú; djádj-in-o-m. In the forms of 'who' and 'what' the $\sqrt{-M-}$ functions as a 'personalizer' or 'identifier' and appears also instead of the other two genders, as well as for plural.)

$\sqrt{-i-}$: with $\sqrt{-In}$ in all adj. (except the pron.adj. 'that', 'all'); 3 pers.pr. (nóv-0-i-m '[with the] new [one]', trjéty-y-i-m, djádj-in-i-m, ét-j-i-m; y-i-m.)

$\sqrt{-e-}$: with $\sqrt{-In}$ in 'that', 'all'; 'who', 'what'. (t-j-é-m, v(je)sj-j-é-m = /fsjém/; k-é-m, č-é-m.)

$\sqrt{-0-}$: with $\sqrt{-N}$ in *y-*, *ov-*, pron.adj., 3 pers.pr. (trjéty-(i)y-0-0, djádj-in-0-0, od(ji)n-0-0, ón-0-0.) — with $\sqrt{-G}$, $\sqrt{-D}$ in *ov-*adj. (djádj-in-0-a, djádj-in-0-u.) — with $\sqrt{-L}$ in listed *ov-*adj. ('proper names'). (pjotr-óv-0-je.) — alone in short forms of reg.adj. (nóv-0-.)

In the numeral 'both' there occurs in the cases other than N the syllable /oy/: G, L /abóyix/, D /abóyim/, In /abóyimji/; these forms are restricted to use with nouns that can be identified as M or Nt when non-plural. It could be argued that there is here an instance of $\sqrt{-M-}$, as $\sqrt{-oy-}$, before and in the presence of $\sqrt{-Pl-}$, as $\sqrt{-i-}$; see §2.4 for the discussion of the corresponding feminine forms and the analysis that we choose.

2.3. The neuter morpheme $\sqrt{-Nt-}$ has three allomorphs: $\sqrt{-oy-}$, $\sqrt{-0-}$, $\sqrt{-o-}$. It occurs only with $\sqrt{-N}$ and in short forms of adjs., its expected place elsewhere being taken by allomorphs of $\sqrt{-M-}$ (i.e. M and Nt 'are the same' except in the N). The distribution is:

$\sqrt{-oy-}$: in reg.adj. (nóv-0-oy-o.)

$\sqrt{-0-}$: in other than reg.adj., and in 3 pers.pr. (trjéty-y-0-o, djádj-in-0-o, ét(o)t-0-0-o = /éta/ 'this', on-0-ó 'it'.)

$\sqrt{-o-}$: alone in short forms of reg.adj. (nóv-o-.)

2.4. The feminine morpheme $\sqrt{-F-}$ has seven allomorphs: $\sqrt{-ay-}$, $\sqrt{-uy-}$, $\sqrt{-oy-}$, $\sqrt{-0-}$, $\sqrt{-ey-}$, $\sqrt{-je-}$, $\sqrt{-a-}$. Their distribution is:

$\sqrt{-ay-}$: with $\sqrt{-N}$ in reg.adj. (nóv-0-ay-a 'new'.)

$\sqrt{-uy-}$: with $\sqrt{-A}$ in reg.adj. (nóv-0-uy-u.)

$\sqrt{-oy-}$: with $\sqrt{-G}$, $\sqrt{-L}$, $\sqrt{-D}$, $\sqrt{-In}$ in all adj. (except listed pron.adj.). (nóv-0-oy-0 for G, L, D; trjéty-y-oy-0, djádj-in-oy-0, t(o)t-0-óy-0 = /tóy/; nóv-0-oy-y(u), trjéty-y-oy-y(u), djádj-in-oy-y(u), t(o)t-0-óy-y(u); in the In forms just listed, there is automatic morphophonemic coalescence of the two y's into /y/, so that phonemically the forms are, after the optional loss of (u), identical with the other three cases.)

$\sqrt{-0-}$: with $\sqrt{-N}$ and $\sqrt{-A}$ in *y-*, *ov-*, and pron.adj. (trjéty-y-0-a, trjéty-y-0-u, djádj-in-0-a, moy-0-0-ú.)

√-ey-: with √G, √L, √D, √-In in listed pron.adj. ('my', 'thy', 'one's own', 'whose', 'all'), and in 3 pers.pr. (except with √-G). (moy-0-éy-0, tvoy-0-éy-0, svoy-0-éy-0, y-éy-y(u).) This allomorph is actually a morphophonemically regular (though not automatic) variant of √-oy-: it occurs in the same kind of distribution, and might be subsumed under √-oy- by stating the morphophonemic conditions; but for the present purpose it seems simpler to list it separately.

√-je-: with √-N in '2', 'both'. (dv-jé-0, ób-je-0.) These forms must be considered F, not Pl. We must say that for these two numerals the forms containing √-F occur in concord with nouns identifiable from their singulars as F, while with other nouns the forms used are dv-á-0 and ób-a-0, containing √-Pl.

√-a-: alone in short forms of reg.adj. (nóv-a-.)

The G, L, D, In of 'both', in concord with nouns identifiable as F, are /abjéyix/, etc., corresponding to the non-F /abóyix/, etc., noted in §2.3. If the latter are to be analyzed, as suggested, as containing both √-M- and √-Pl-, then the forms here contain both √-F- and √-Pl-, and there is an additional allomorph of √-F-, namely, √-jey-. Since, however, these form of 'both' are unique, and since Russian has many paradigms where the N forms show different stems from the other cases, we prefer to say that 'both' has the stems √ob-, used with √-Pl- and √-F- before √-N, and the two stems √oboy- and √objey- both used before √-Pl- (as √-i-), the latter in concords where N /óbji/ would occur.

It may be pointed out here that the allomorphs of the gender morphemes show predominantly the forms -Vy- and -V-. These may be considered morphophonemically regular alternants of each other: the -y- appears when a vowel or zero follows in the case suffix, and is absent before a consonant; the -V- is basically -o- in all three, but they are distinguished as separate morphemes by the other allomorphs that occur. These other allomorphs are, in terms of this formulation, the irregular ones.

2.5. The possibility of treating the gender, number, and case morphemes as we have done here—i.e. as only thirteen different morphemes, with allomorphs—depends upon the possibility of establishing determining criteria in the stems. For nouns this is done in terms of the classification by gender and declension-class. Adjectives might then be treated simply as additional types of stems, which are inflected for gender rather than classified by it. When, however, the morphemes attached to noun-stems have been set up as above, and the attempt is made to follow the same pattern for the morphemes attached to adjectives, it becomes clear that at least some of the forms contain additional morphemic material. It has therefore seemed best to state the conditions for the presence of the various allomorphs found with adjectives in terms of a morpheme of adjective-declension, √-Adj-. (Such a morpheme also provides a simpler way of accounting for the isolated forms, adjectivally declined in one gender only, that function as nouns—e.g. /partnóy/ 'tailor'. The presence of the √-Adj- provides a formal basis for the declension as an adj., and necessarily calls for a gender-morpheme.) The √-Adj- is here considered to be one morpheme only, with the six allomorphs √-iy-, √-(i)y-, √-in- ~ √-ov-, √-j-, √-0-. It might be argued that this is a stem-forming morpheme; but it is not 'derivational' like certain suffixes found in adjectives. Accordingly, since we are here dealing with inflection

and Russian derivation has still to be examined, we shall regard the morpheme in question as inflectional (pre-final), at least for the present. The allomorphs of this morpheme are, in fact, determined by the listing of adjective-stems in the four types that we have been naming in our treatment so far. Their distribution is:

✓-iy-: with ✓-Pl- (as ✓-a-) and ✓-N (as ✓-0) in reg.adj. (nóv-iy-a-0.)

✓-(i)y-: in all *y*-adjs. (trjétj-(i)y-0-0, trjétj-(i)y-0-o, trjétj-(i)y-i-x, etc.)

✓-in- ~ ✓-ov-: in all *ov*-adjs.; the two forms are listed as alternants because they are mutually exclusive. (djádj-in-0-0, pjotr-óv-0-0).

✓-j-: with ✓-M- (as ✓-i- and ✓-e-) in In (as ✓-m), and with ✓-Pl- (as ✓-i- and ✓-e-) in all the cases, in pron.adj. (ét-j-i-m, t-j-é-m; ét-j-i-0, ét-j-i-x, ét-j-i-m, ét-j-i-mji, t-j-é, etc.)

✓-0-: in all forms not listed for the other allomorphs, namely: for reg.adj. all case forms Sg and Pl, except Pl N; for pron.adj. all M (Sg) except In, the Nt (N), all F forms.

In the M N of reg.adj. the analysis of the /iy/ variant could be different from that given above. We have treated the /iy/ as a variant of M ✓-oy-, and analyzed the ending as having Adj. ✓-0-, M ✓-oy-, N ✓-0. In view of the setting up of Adj. ✓-iy-, however, in the Pl N ✓-iy-a-0, we could analyze the M N as ✓-iy-0-0; this would make all adj. M N parallel in having two terminal zeros—one M and one N. There is really little to choose between the two possibilities. We prefer the one that makes /iy/ a variant of M ✓-oy-.

In the short forms of reg.adj., the only overt inflection consists of the endings -0-, -o-, -a-, -i-; these have been analyzed above as the gender and plural morphemes alone, without preceding ✓-Adj- or following case morphemes. This seems the easiest and most efficient analysis, and can be applied also, as already mentioned, to the past forms of verbs.

The analysis of the short forms raises the question whether we cannot leave out the ✓-Adj- as such altogether, and simply classify adjectives as having stems in /y/, /in/ ~ /ov/, /C/ ~ /Cj/, and /C/ (reg.adj.), with different allomorphs of the gender and plural morphemes determined by the stem (as for plural in nouns). Such an analysis, however, would be unsatisfactory because of the difficulty of then analyzing the reg.adj. Pl /iya/. If this involves a Pl ✓-iy-, the final /a/ must be the ✓-N; in that case, other instances of Pl N /a/ ought to be analyzed in the same way. This would bring with it a whole series of other rearrangements in what is otherwise a rather neat analytical picture. Moreover, the distribution of the stem alternants in /C/ and /Cj/ in pron.adj. is not paralleled elsewhere in the language except possibly in the 3 pers.pr. All in all, the analysis with a morpheme of adjective-declension seems best, even though this calls for so many instances of ✓-0-, and, in the pron.adj. M N, for three zeros in a row.

3. The distribution of stress in Russian is such that the stress-patterns function as word-completing morphemes. When, within a paradigm of inflection, there are variations in stress, we can speak of a paradigm of superfixes. The variations found, with a few exceptions, are limited to alternation between 'stem stress'—primary stress on some syllable of the stem, and 'ending stress'—

primary stress on the vowel (or the first of two vowels) of the desinence. It is easy to indicate these alternations by a formula such as AB, where A means stem stress (say) in the Sg, and B ending stress in the Pl. For nouns, this has been done several times in the past; for other entitives the statements have not so far been made.

3.1. The 'stress patterns' of nouns have most recently been treated by Olmsted.³ We accept his statements, with one emendation, and summarize them here, with explicit restatement of the actual superfixes involved.

Three patterns each are found in Sg and Pl: A—primary stress on the same syllable of the stem throughout; B—primary stress on the vowel (or first of two vowels) of the desinence, or on the movable vowel before a zero-desinence, or on the last stem vowel before a zero-desinence; C—like B, but like A for the Sg A (and N, if they are alike) and for the Pl N. For a full indication of the paradigm, two letters must be used: one for the Sg and one for the Pl. If there is more than one stem vowel, we distinguish between A₁ for stress on the last vowel, A₂ for stress on the second-last, etc. (this is not Olmsted's numbering). The patterns, with Olmsted's examples, are: AA škóla 'school', AB slóvo 'word' (Pl slová), AC četvjortj 'quarter' (Pl N četvjortji, G četvjortjéy, etc.); BA ljícó 'face' (Pl ljíca), BB nóž (knife) (G nožá, Pl N noží); CA zorjá 'dawn' (A zórju), CB úgol 'corner' (G uglá, Pl N uglí), CC ruká 'hand' (A rúku, Pl N rúki, Pl L rukáx). Also cited is A₂A ózjoro 'lake' (Pl ozjóra). Olmsted says that BC does not occur, but at least some speakers have BC zjemljá 'earth, land' (A zjemljú, Pl N zjémli).

The Pt case follows the pattern of the Sg G for any noun where it occurs. The L₂ case always has stress B. The Vo has been illustrated for the forms where it occurs (§1.9); one of the instances shows stress A counter to the rest of the Sg ('Father'), one has A₁ as against A₂ ('Lord'), the other three follow the Sg pattern.

Various superfixes are involved in these patterns. When a noun stem has only one vowel, pattern A means $\sqrt{\text{v}}$ for the forms with one-vowel endings (škól-a), $\sqrt{\text{v}\text{v}}$ for those with two-vowel endings (Sg In I optionally, Pl In: škól-oy(u), škól-amji), and $\sqrt{\text{v}}$ for those with zero endings (Sg N II M, Pl G dóm-0, škól-0). If the noun stem has two vowels, A₁ means $\sqrt{\text{v}\text{v}}$, $\sqrt{\text{v}\text{v}\text{v}}$, or $\sqrt{\text{v}\text{v}\text{v}\text{v}}$, while A₂ means $\sqrt{\text{v}\text{v}}$, $\sqrt{\text{v}\text{v}\text{v}}$, or $\sqrt{\text{v}\text{v}\text{v}\text{v}}$. In longer stems there are correspondingly longer superfixes: $\sqrt{\text{v}\text{v}\text{v}\text{v}}$, $\sqrt{\text{v}\text{v}\text{v}\text{v}\text{v}}$, etc. Pattern B always means $\sqrt{\text{v}\text{v}\text{v}}$ or $\sqrt{\text{v}\text{v}\text{v}\text{v}}$.

3.2. Stress patterns for adjectives will be stated first in general terms. Regular adjectives always have the primary stress on the same syllable throughout the paradigm, except that some or all of the short forms may be different. The same statement applies to *y*- and *ov*-adjectives. The primary stress is either on the stem (A, A₁, A₂, etc.) or on the suffix (the adjective-declension or the gender-number morpheme: B). Short forms of reg.adj. have combinations of A and B, detailed below. In pron.adj. the prevailing pattern is one of stress on the last vowel except the last vowel of Pl In). Specific delimitations follow.

Regular adjectives with stress A are illustrated by /nóviy/, with such forms

³ David L. Olmsted, The morphophonemics of Russian noun inflection, *SIL* 9.1-6, esp. §§2.1-5 (1951).

as /nóvaya/, /nóvimji/. A₁ is illustrated by /glubókiy/ 'deep', A₂ by /grjéčiskiy/. The actual superfixes for A are $\sqrt{\text{zv}}$ and $\sqrt{\text{zvv}}$; for A₁, $\sqrt{\text{zv}}$ and $\sqrt{\text{zvv}}$; for A₂, $\sqrt{\text{zv}}$ and $\sqrt{\text{zvv}}$. Stress B is found in such adjectives as /baljšóy/ 'big', giving superfixes $\sqrt{\text{...z}}$ and $\sqrt{\text{...zv}}$ as in /baljšáya/.

Illustrations for *ov*-adjectives: /djádjín djádjína djádjínimji/; /pjitróf pjitróva pjitróvimji/—superfixes $\sqrt{\text{zv}}$, $\sqrt{\text{zvv}}$, $\sqrt{\text{zvvv}}$ for A₂, $\sqrt{\text{...z}}$, $\sqrt{\text{...zv}}$, $\sqrt{\text{...zvv}}$ for A₁; there are B types only for -in-: /iljyín/ (family name). The *y*-adjectives have only A patterns: /trjéttij trjéttijimji/— $\sqrt{\text{zv}}$, $\sqrt{\text{zvv}}$.

Short forms are of these types: all four forms with primary stress on the stem; F with primary stress on the suffix; F and Pl with primary stress on the suffix; all except M with primary stress on the suffix and M with primary stress on the last vowel (including the movable vowel). These may be designated respectively A, B_F, B_{Pl}, B. Examples: A /bagát/ 'rich', /bagáta/, /bagáta/, /bagáti/; B_F /sláp/ 'weak', /slába/, /slabá/, /slábi/; B_{Pl} /pólan/ 'full', /pólna/, /palná/, /palní/; B /xaróš/ 'good', /xarašó/, /xarašá/, /xaraší/.

In pron.adj. the stress pattern (of all except 'this', 'our', 'your', and 'self'), primary stress always on the last vowel except in Pl In, may be called B+. The superfix is $\sqrt{\text{...z}}$, Pl In $\sqrt{\text{...zv}}$: /adjín/ 'one', /adná/, /adnavó/, /adnjímji/. In 'self' there is the same pattern except that the Pl N is /sámji/; this may be called C+. In 'this', 'our', 'your' the pattern is A.

3.3. The pronouns all have primary stress on the suffix vowels when these occur, so that the pattern can be called B. In /yá/, /tí/, /ón/ our analysis sets up $\sqrt{\text{v}}$ -0 of the N, so that the stem stress is the regularly expected result in pattern B. In the 3 pers.pr. the Sg has B+, as in /yivó/, /yimú/. This is also true of 'who' and 'what': /kavó/, /čimú/.

3.4. Most of the numerals show suffix stress, pattern B: /dvá dvúx trjimjá pjítj šístjyú stó/. In /čítirji čítirjóx/, /vósjim(j) vasmj/, /djévjitj djívjitj/, /djésitj djisjítj/, /sórak saraká/, the pattern is C. The compounded numerals show A in the teens and C in '20' and '30'. The higher tens and the higher hundreds are inflected as phrases, each part separately, and need not concern us here.

4. The patterns of allomorph distribution and the morpheme sets described above are intended to cover the whole of Russian declension, except possibly for unusual irregularities in superfix alternation and occasional obsolescent or petrified forms. Not covered, however, is the inflection of such items as /póldjinj/ 'noon', /paltará/ 'one and a half' /pjídjisját/ 'fifty', etc.; these, though syntactically units, still show their structure as phrases rather than words. The declensional suffixes in such phrases are easily enough identified in terms of our analysis; the principal characteristics of the phrases are their superfix paradigms, which must be treated along with other phrasal superfixes.

DRAVIDIAN KINSHIP TERMS

M. B. EMENEAU

University of California

[Proto-Dravidian had a class of kinship nouns which occurred only in the possessed construction (inalienably possessed); this probably was a syntactic rather than a morphological construction. The personal and reflexive pronouns which occurred as attributes in this construction were only the plural ones; distinction of number in the possessor was not indicated, and could be gathered only from the context. The evidence is drawn from Old Tamil, Kota, Gondi, Kolami, Kuwi, and Kurukh.¹]

Srīgurave namaḥ.

1. It has long been noticed² that modern Tamil *tampi* 'younger brother or male parallel cousin' and *taṅkai* 'younger sister or female parallel cousin' are each the last survivor of an earlier set of four forms. The Old Tamil literature shows: *empi* 'my younger brother', *numpi*, *umpi* 'your younger brother', *tampi* '(his or her own) younger brother'; *eṅkai* 'my younger sister', *nunkai*, *unkai* 'your younger sister', *taṅkai* '(his or her own) younger sister'.³ The old literature

¹ The bibliography is as in *Lg.* 21.184 fn. 1, with additions as given further on in this note.

The sigilla for the languages are as follows: Dr. = Dravidian, PDr. = Proto-Dravidian, Ta. = Tamil, Ma. = Malayālam (H. Gundert, *A Malayalam and English dictionary*, Mangalore, 1872), Ka. = Old Kannaḍa (F. Kittel, *A Kannaḍa-English dictionary*, Mangalore, 1894), Koḍ. = Koḍagu (Coorg), Tu. = Tulu, Te. = Telugu (P. Sankaranarayana, *A Telugu-English dictionary*, Madras, 1927), Kol. = Kolami (usually my field material; SR = P. Setumadhava Rao, *A grammar of the Kolami language*, Hyderabad, 1950; Kin. = dialect of Kinwaṭ, Adilabad, Nk. = Naiki; Pa. = Parji; Oll. = Ollari; Go. = Gondi (without further sigillum, C. G. Chenevix Trench, *Grammar of Gondi*, Madras, 1919, 1921; M. = A. N. Mitchell, *A grammar of Maria Gondi as spoken by the Bison Horn or Dandami Marias of Bastar State*, Jagdalpur, 1942; L. = A. A. Lind, *A manual of Madiā*, Kedgaon, 1913 [not available to me, but quoted sometimes from Burrow]; W. = H. D. Williamson, *Gondi grammar and vocabulary*, London, n.d. [1890]); Kur. = Kurukh (Oraon); Malt. = Malto; Br. = Brahui. Kota, Toda, Kui, and Kuwi are not abbreviated; Kuwi (F) = A. G. Fitzgerald, *Kuvinga Bassa, The Khond language as spoken by the Parjas ... of the Madras Presidency*, Calcutta, 1913; Kuwi (S) = F. V. P. Schulze, *Vocabulary of Kuwi-Kond language*, Madras, 1913. IA = Indo-Aryan. Turner = (Sir) Ralph Lilley Turner, *A comparative and etymological dictionary of the Nepali language*, London, 1931.

I am indebted to Professor T. Burrow of Oxford University and Sri Sudhibhushan Bhattacharya of the Indian Museum, Calcutta, for allowing me to use items collected by them during their field trips in Bastar and in the Adilabad district of Hyderabad; the languages quoted on their authority are the Kinwaṭ dialect of Kolami, Naiki, Parji, and Ollari. My work has benefited much from discussion with Burrow.

A summary account of the phonetic symbols used will be found in *BSOAS* 14.98 fn. 1.

² Especially Robert Caldwell, *A comparative grammar of the Dravidian languages*³ 397-400 (1913); Julien Vinson, *Manuel de la langue tamoule* 69-70 (Paris, 1903). Caldwell undoubtedly had examined the material in his 2d edition of 1875 and perhaps even in the 1st edition of 1856; I have no easy access to either of these editions. Most recently P. Meile has discussed some at least of these forms in a lecture which is reported in outline in *BSL* 46.1.xiii (1950).

³ The *Tamil lexicon* glosses *empi* as *eṅ tampi*, *eṅkai* as *eṅ taṅkai*, etc.

contains a similar set for 'father': *entai*, *nuntai/untai*, *tantai*, and a partial set for 'elder brother': *emmun*, *nummun*, *tammun*; no forms from these two sets survive into the modern language. The meanings given are those of the *Tamil lexicon*, except that I have added '(his or her own)' for the *t*-forms. The *Lexicon* records also for *entai* the meaning 'our father', and it is to be suspected that the meanings should be given more exactly as: '...of me or us (exclusive)', '...of you (sg. or pl.)', '(his, her, or their own)...'. Of the 2d personal forms those with *n-* are the older and are displaced historically by those without *n-*.

2. Analysis of these two-membered forms (comparison with other constructions yields information on where to make the cut) gives the following morphemes (*N* = nasal homorganic with following stop or nasal): *eN-* 'my or our (exclusive)', *nuN-/uN-* 'of you (sg. or pl.)', *taN-* 'his, her, or their own'; *-pi* 'younger brother or male parallel cousin', *-kai* 'younger sister or female parallel cousin', *-tai* 'father', *-mun* 'elder brother'.

3. The two-membered forms when compared with similar forms in Tamil are clearly examples of attribute-head constructions. The chief difference in the present instances is that *-pi*, *-tai*, and *-mun* occur only as bound forms in these constructions. *kai* 'younger sister' is recorded by the *Tamil lexicon* as a free form occurring in *Piṅkalanikaṇṭu*, a dictionary, as is also *kaiyai* with the same meaning; Caldwell (400) says that in Tamil poetry *kaiyai* is used in the sense 'younger sister', but he gives no references. *-mun* 'elder brother' is homophonous with *mun* 'in front, previous, prior; that which is first or chief' and is in all probability derived from it ('he who is prior to me' etc.). *-pi* and *-tai* cannot be further analyzed,⁴ and must be taken as bound forms occurring in no other constructions than those with the preceding pronominal forms.

4. The prior members in the constructions, viz. *eN-*, *nuN-/uN-*, *taN-*, which have pronominal possessive meaning, are to be identified with the oblique stems of plural personal pronouns: 1pl. exclusive *em-*, 2pl. *num-/um-*, reflexive pl. *tam-*, which are also used as independent words with possessive meaning; they are not to be identified with the singular personal pronoun forms: 1sg. *eṇ-*, 2sg. *nin-/nun-/un-*, reflexive sg. *tan-*. There are two reasons: (1) the older 2d personal form (*nuN-*) has the vowel *u* of the 2pl. rather than *i* of the older 2sg.; (2) the external sandhi rule is that *m* alternates with the nasal homorganic with a following stop, while *n* remains unchanged. Further confirmation is provided by the following literary material quoted by the *Tamil lexicon*: *num aiyan* 'your elder brother', *tam aiyan* '[his] elder brother', *tam appan* '[his] father', *tam akkai* '[his] elder sister'; in these constructions *num* and *tam* show the basic *m*-alternants of the plural pronouns before vowels. One would expect doubling of the *m* in all these instances.⁵ I have found no explanation of the failure to follow

⁴ Caldwell (399) suggested that *-pi* was derived from *piṇ* 'back, rear part, that which is subsequent in time', i.e. 'he who is subsequent to me', etc. He saw the difficulty, viz. the lack of final *-n* in the kinship words. The presence of *piṇ* 'younger brother' in the old glossary *Tivākaram* (Cēntaṇ-tivākaram) is probably nothing but an attempt at an etymology.

⁵ Final short *l*, *n*, *l*, *ṇ*, *y* (and *m*) of original monosyllabic words with short vowels become lengthened when immediately followed by vowels in sandhi contexts.' So L. V. Ramaswami Aiyar, *Quarterly journal of the Mythic Society* 26.173 (1936). He notes that *em-*, *num-*, *tam-* have such doubling.

the rule. However, *emm anai* 'our mother' and *tamm anai* '[his] mother' follow the rule, though it is notable that *anai* seems to occur only in these two constructions instead of *annai* 'mother' (see Etymologies A, no. 2).

5. Old Tamil, then, has some kinship terms which only occur bound with pronominal possessives. We shall use for this feature a term that has been used in describing other languages, viz. INALIENABLE POSSESSION. The pronominal possessives in Old Tamil are always plural in form, there being no indication of a difference of number in the possessor (see §1 for the slight uncertainty on this point). Modern Tamil has lost this construction and uses the reflexive form throughout all persons, preceded by either singular or plural possessive forms as the context may demand; e.g. *en tampi* 'my younger brother', *un tampi* 'your younger brother', *enkaḷ tampi* 'our younger brother'.⁶

6. The forms possessed by the 1st person exclusive plural *em-*, *empi*, *enkai*, *entai*, *emmun*, are presumably those that would be statistically most common in actual kinship usage (as opposed to the 1st person inclusive plural), i.e. they would be used by a person identifying his kin to someone outside his family or to someone within the family but in a different kin relationship to the person referred to from that of the speaker. We should expect also to find the same morphemes possessed by the 1st person inclusive plural *nam-*, the forms being *nampi*, *naṅkai*, *nantai*, *nammun*. Of these, the last is not recorded. Caldwell and Vinson mention *nantai* as occurring in the poets with the meaning 'our father' (not distinguished from *entai* in their report); the *Tamil lexicon* does not list the word. *nampi* has the following meanings: 'the elite among men, a perfect soul, the Supreme Being, the title of officiating temple-priests, a term of endearment', and *naṅkai* the following: 'lady, woman of quality or distinction, son's wife, elder brother's wife'. If these are the relationship terms that we (as well as Caldwell and Vinson) have assumed, real kinship statuses have been displaced by supposititious ones which would be the same for the speaker and the one spoken to; consequently, the inclusive 1st plural pronoun would be in order.

7. Almost all the Tamil kinship forms that are relevant to the problem of possessed forms have been given.

There remain three sets of doublets: *annai* 'mother, elder sister', *tannai* 'mother, elder sister, elder brother'; *ammai*, *tammai* (so Vinson, 69) 'mother'; *attai* 'father's sister, mother-in-law, lady', *tattai* 'elder sister'. It was already suggested by Caldwell that the *t*-forms show *ta-*, an allomorph of the plural reflexive pronoun *tam-*. It is not at all clear why he postulated *ta-* rather than *t-*. The sandhi rules in Tamil for final vowels (other than *-u*), when a vowel begins the next morpheme, in general require a glide consonant between the vowels, and not elision or contraction. However, there are instances, both in Old Tamil and in the modern colloquial, of elision of the first of the two vowels.⁷

To anticipate the comparative evidence, the Central Indian languages Kui (§11) and especially Gondi (§13) have *t*-forms which present the same problem. Kota of the Nilgiri Hills (adjacent to the Tamil area) has clear instances of *ta*

⁶ Note that *tan*, singular reflexive, occurs in similar constructions (Vinson, 83): *en tan karpu* 'my chastity', *anuman tan uḷal* 'Hanuman's body'.

⁷ L. V. Ramaswami Aiyar, *QJMS* 26.94 f.

preceding consonants in parallel constructions, but in these there has been simplification of clusters made up of nasal and homorganic stop, by loss of the nasal (§12). If *ta-* is postulated for the Tamil forms, the allomorph has a somewhat more usual form for the first member of a compound than if *t-* is postulated. However, *t-* could be taken to be analogical to the other forms with *t-* (e.g. *tampi*).

It should be said that *tannai*, *tammai*, *tattai* are more likely than *tam appan*, *tamm anai*, etc. to be contemporary in formation with *tampai*, *tanikai*, etc., since they are less regular than *tam appan*, etc. The latter look as if they might be analogical formations (from **tappan*, *tannai*, etc.) on the basis of constructions in which *tam* preceded a noun that was not a kinship term.

8. Another set of Tamil words requires discussion: *āy*, *yāy*, *ñāy*, *tāy* 'mother'. So the *Lexicon* gives the meaning of all four forms. *āy* and *tāy* must be related in the same way as the pairs of forms discussed in §7. Jules Bloch⁸ specifies that *yāy* means 'my mother' and *ñāy* 'your mother' (for the latter Meile, in the place referred to in n. 2, gives a reference to the old text Kalittokai 107, 26). This being so, it is tempting to find in *yāy* some representation of *em-*, perhaps merely *e-*, and in *ñāy* some representation of *num-*. Without parallels, however, one can go no further, and it is to be noted that *y-* appears again in Gondi *yāyāl* 'mother'; Burrow (BSOAS 11.599) has connected the alternation of *yā-* and *ā-* with that seen in Ta. *yāḷu*, *āḷu* 'goat', *yāru*, *āru* 'river', etc. The form with *ñ-* is highly peculiar; Bloch states '*ñ-* de **nī*', i.e. he takes *ñ-* to be somehow derived from *nī* 'you (sg.)'.

9. Tamil has also in old texts *tallai* 'mother', which seems to be Proto-Dravidian because of its wide range of occurrence (Ta.-Ma., Te., Pa., Kui-Kuwi; see Etymologies A, no. 9). The initial *t-* hardly yields to analysis, but is to be noted. It may be suggested that this noun belongs to the sets which are under discussion; there is, however, no evidence at all that there is anything more than accident involved in the similarity of initials.

10. Tamil *takappan* 'father' has been discussed by Caldwell and Vinson as containing *tam*. This analysis can be achieved only by positing a highly unusual shift $m > k$. Rather, the first member of the construction is *tak(u)*, the adjectivally used stem (Vinson's 'participe abrégé', 126 f.) of the verb *taku* 'be fit, appropriate, suitable, worthy, adequate, be excellent', i.e. 'excellent father'. The construction is entirely parallel to that of *aḷu-kalīru* 'a killing elephant',⁹ even to the detail that *taku* (*takuv-*, *takk-/takunt-*) and *aḷu* (*aḷuv-*, *aḷḷ-*) 'smite, kill, destroy, conquer' belong to the same verb class (*taku* has an alternative membership in another class as well); for these two verbs, see Etymologies A, nos. 11 and 12.

11. The forms that the other Dravidian languages have cognate with Tamil *tampi*, *tanikai*, *tāy*, and *tantai* are given in the Etymologies A, nos. 5-8. They need not be examined in detail. It must be noted that the reflexively possessed forms are the only ones that occur in the other languages, as is true in modern

⁸ *Structure grammaticale des langues dravidiennes* 24 (Paris, 1946).

⁹ Quoted from *Puranāṇūru* 69 by P. S. Subrahmanya Sastri, *Tolkāppiyam-Collatikāram* (Annamalai University Tamil series, No. 9; 1945) in his commentary on sūtra 415.

Tamil for *tampi*, *taṅkai*, and *tāy*. Tamil is the only one of the languages that provides us with a history of the forms, and in Tamil it is evident that the extension of the use of the *t*-forms has taken place within the history of the language itself. This would seem to be an excellent instance of what Sapir called 'drift'. The languages involved are in general the southern ones (including Tuḷu and Telugu, but not including the Nilgiri languages Toda and Kota), plus a scattering of the central ones, but not Kurukh-Malto (or Brahui, for which evidence is missing from all these etymologies). That the phenomenon cannot be due to borrowing over the whole of the large, almost unbroken southern and central area seems to be guaranteed by such facts as Kui's *tangi* 'younger sister' contrasting with the quite different words used by Telugu, Parji, and Gondi (the words are clearly related: Te. *celli*, *celliya*, *celliyalu*, *celle*, *cellelu* 'younger sister'; Parji *cāla* 'sister'; Go. *sēlār*, *sēlār*, [W] *selār* 'younger sister', [M] *helār* 'sister') and Kui's *tambesa* 'younger brother' with a different suffixal formation from that of Telugu *tammūdu* and Gondi *tāmmūr*. The words for 'mother' (Ta. *taḷḷai*) may be an example of the same 'drift', or may have acquired a fixed form with *t* in Proto-Dravidian (if indeed the *t*- is of the same reflexive origin); Old Tamil here fails to give any relevant evidence.

The Kui forms *tangi* and *tambesa* have beside them *aṅgi* and *ambesa*. Winfield fails to describe any difference in usage, and the forms without *t*- remain a problem.

12. In the grammatical sketch in *Kota texts* (1.24, §38), it was stated that the singular personal and reflexive pronouns have allomorphs *e*, *ni/di*, *ta* which occur as attributes before nouns that begin with a consonant and denote relatives of the person denoted by the attribute. The two forms of the 2d personal pronoun were stated to occur in the following distribution: *ni* if the relationship term begins with a nasal, *di* if it begins with a stop. Actually, *ni* occurs before *ma·mn* 'father-in-law, maternal uncle, father's sister's husband', *ma·ym* 'mother-in-law, paternal aunt, mother's brother's wife', *meyṇ* 'son', *mo·l* 'daughter', *na·tu·ny* 'woman's female cross-cousin or sister-in-law', and also before *jaḍukn* 'wife's sister's husband (if the men are of different clans)'; *di* occurs before *peḍ* 'wife', *kara·l* 'younger brother or male parallel cousin', *kara·c* 'younger sister or female parallel cousin', *doḍa·c av* 'mother's elder sister', *pe·r i·ṇ* 'grandfather', *pe·r av* 'grandmother', with *k*, *p* > *g*, *b*. The informant gave *e mo·l* with the meanings 'my daughter, daughter of us (exclusive)'; occurrences in texts guarantee *ta mo·l* 'his, her, their daughter', *ta beḍ* 'his, their wife', *ta ga·ra·c* 'his, their younger sister'. In these examples, that is, and undoubtedly in all the rest as well, the pronominal forms do not show the number of the possessor or possessors.

The allomorphs *e*, *ni*, *ta* before the words beginning with *m*- are to be taken as allomorphs of the plural pronominal forms 1st exclusive *em*-, 2d *nim*-, reflexive *tam*- (rather than of the singular forms *en*-, *nin*-, *tan*-), since quite regularly there is morphophonemic replacement of a cluster of identical consonants by a single consonant (*Kota texts* 1.17, §10.1a). We should then take all the instances of the allomorphs *e*, *ni/di*, *ta* in the same way. They are allomorphs of *em*-, *nim*-, *tam*-. The morphophonemic alternations involved occur elsewhere in the

language: $mp > b$, (mt , $md >$) $nd > d$, ($mk >$) $\eta g > g$ (it is to be noted that there are no examples of $np > b$, etc.). For other examples, see Etymologies B.¹⁰ The use of these allomorphs before *na·tu·ny* and *jaḍukn* is unexplained in phonological terms; analogy is probably to be invoked.

In Kota, therefore, the only pronominal possessive forms that occur before certain of the kinship terms are the plural ones; the number of the possessor or possessors is not indicated.

None of the forms with initial *t*- discussed in the treatment of Tamil occurs in Kota, nor has Kota any kinship terms that are inalienably possessed.

13. Gondi has not only *tāngē* 'younger sister' and *tāmmūr* 'younger brother', but also *tānnāl* 'elder brother' (Etymologies A, no. 3) and *tākkā* 'elder sister' (but Lind's Madia has *akkā*; Etymologies A, no. 1). It seems most probable that these latter two words have *t*- by analogy with the forms for the younger siblings.

Trench records that the Gondi words *āpōrāl* 'father' and *āwūl* 'mother' (Etymologies A, no. 4) when possessed show forms *maipō* 'my father', *māipō* 'thy father', *maiwūl* 'my mother', *māwūl* 'thy mother'. Exact analysis is not possible because of bad phonetic recording, but the possessives are allomorphs of the plurals *mā*- 'of us' and *mī*- 'of you (plural)'.

14. Kolami has at least four kinship terms which always occur syntactically possessed: *ba·n* 'father', *ay* (? *a·y*) 'mother', *bai* 'sister', *ta·k* 'father'. The last occurs in such constructions as *anne* (? *ane*) *ta·k* 'my father', *amne ta·k* 'his father', *avre ta·k* 'their father', *ma·sale ta·k* 'wife's father', *avre ta·k ammaner* 'their parents' (Naiki has *tāk* 'father' and *amma tāk* 'parents'). The other three, as well as some other kinship terms (which also occur freely without a possessor), have only plural pronouns as pronominal possessive attributes; the number of the possessor is not indicated. Examples: *am ba·n* 'my father, our (exclusive) father', *im ba·n* 'your (sg. or pl.) father', *tam ba·n* 'his own father, her own father, their own father', *avr ba·n* 'his father, her father, their father' (*avr* 'they'), *avr ay* 'his mother, her mother, their mother', *am bai* 'my sister, our (exclusive) sister', *avr vanna ini avr da·dak* 'her elder brother's wife and her elder brother'.

15. In Kuwi as described by Fitzgerald, *abba* 'father' and *īya* 'mother' have possessed forms *māba* 'my father', *mība* 'your father', *maiya* 'my mother', *mīya* 'your mother'. The pronominal prior elements are allomorphs of the plural pronouns.

Schulze records much the same forms for these nouns and similar forms for 'grandfather' and 'wife': *mā akku* 'my grandfather', *mīai* 'your wife' (but *nāai* 'my wife' with a singular pronominal form), *māija māba* ($j = y$) 'my mother and my father', *mība mīja* ($j = y$) 'your father and your mother'. It is to be noted that he gives *nā tanzi* ($z = j$) 'my father' and *nā talli* 'my mother', as well as other kinship terms, with singular pronominal forms. He also translates

¹⁰ This type of simplification of homorganic nasal and stop by loss of the nasal is chronologically prior to the simplification by loss of the stop that was stated in *Kota texts* 1.18, §11.9.

The alternation of *ni* and *di* is to be classed as partial assimilation of nasal (*n*-) to the stop-type of production that begins the following words (*b*, *d*, *g*).

māba and *māija* as 'our father' and 'our mother'; the number of the possessor is not indicated.

16. In Kurukh a number of kinship terms are always possessed: *-bas* 'father', *-yō* 'mother', *-das* 'son', *-dā* 'daughter', *-ris/-dis* 'younger brother', *-rī/-dī* 'younger sister'; e.g. *iṅgyō* 'my mother', *naṅgdī* 'our (inclusive) younger sister'. Since *bā* 'father' and *yō* 'mother' occur as single words as vocatives, it is fairly certain that all these combinations should be taken as two words each, the kinship term being syntactically inalienably possessed, i.e. *iṅg yō*, *naṅg dī*, etc.

One of these terms, *-bas* 'father', is only preceded by the plural pronouns and the number of the possessor is not indicated: *embas* 'my father, our (exclusive) father', *nimbis* 'your (sg. or pl.) father', *tambas* 'his, her, their own father'.¹¹

17. In an attempt to find PDr. features in the material just presented, two questions must be considered separately. (1) Is there evidence for inalienable possession of kinship terms as a PDr. feature? (2) Is there evidence that in PDr. kinship terms were preceded by pronominal possessive forms that were only plural, the number of the possessor not being indicated?

18. There are two good lines of evidence that inalienable possession of kinship terms was a PDr. trait. One has been set forth already in §11, viz. that the *t*-forms with which we started our examination, both those of Tamil and those of the other languages, are explicable historically only as relic forms of an older system of inalienably possessed kinship terms such as Old Tamil alone preserves.

The second argument is that Old Tamil, Kolami, and Kurukh all show this feature for some terms. They are independent languages, and so far as the evidence goes they have not been in close enough contact to force us to posit borrowing. Kolami and Kurukh are at present not in contact at all, nor have they enough peculiar resemblances to make it seem probable that they have been in recent contact. Moreover, the details of their systems of kinship terms are strikingly different. A priori, there is nothing against either or both of these languages having been southern languages originally, and even having been either Old Tamil dialects or influenced by Old Tamil dialects. But in fact, they are not closely allied with the southern languages. Kurukh, Malto, and Brahui go together in some details as a northern group. Kolami, Naiki, Parji, and Ollari form a separate central group, different from Telugu, Kui-Kuwi, and Gondi (as will be clear when Burrow has published his Parji material and I my Kolami). Since, therefore, the particular feature now being discussed occurs in three independent languages, we may posit it as a feature of Proto-Dravidian.

Whether inalienable possession of kinship terms was a morphological or a syntactic feature is not so certain, and perhaps depends on each of the languages concerned being described according to the same general system. My description of Kolami treats it as a syntactic feature, for sufficient reason within the total system of Kolami.¹² At least one obvious feature of Kurukh, the presence of

¹¹ There is no question here of assimilation of *ṅg* to *m* before *b*; contrast *eyg bāyṅālas* 'my husband's elder brother'.

¹² The phonemic system of Kolami is simplest if word divisions are so placed that long vowels occur only in the first or in the only syllable of a word. Consequently, *avreta-kammaner* 'their parents' has a word division putting *a-* in the first or only syllable of a word. Com-

two unpossessed vocative forms, looks in the same direction. The *t*-forms of modern Tamil and the other languages are unanalyzable, and the dictionaries that handle the Old Tamil forms treat the feature of inalienable possession as a morphological feature, though I arrive at this statement by putting into explicit terms what seems to be implicit in the practice of writing *empi*, *tantai*, etc. as undivided words. A more explicit analysis of Old Tamil, however, would arrive at the same analysis that I reach for Kolami. If so, then the feature of inalienable possession of kinship terms in the three languages, and so presumably in Proto-Dravidian, is a syntactic feature.

Is it possible to arrive at a list of PDr. inalienably possessed kinship morphemes? Those that are widespread in the various languages are presumably Proto-Dravidian. This includes certainly the morphemes for 'father' (Ta. *tantai*), 'mother' (Ta. *āy*, *tāy*), and 'younger sister' (Ta. *taṅkai*). Because of the possibility that two different constructions are involved in the 'younger brother' words (Etymologies A, no. 8), we cannot be quite so certain about this one; yet Ta.-Ma. (*tampi*) and Kui (*tambesa*) seem to have the same morpheme, and Ka., Kod., Tu., Te., Kol., and Go. clearly have one morpheme in common, and I may be too scrupulous in finding a difficulty here. The morphemes for 'father' seen in Kol. *ba'n*, Kur. *bas*, and Kuwi *māba*, *mība* hardly point to a PDr. morpheme, particularly since Indo-Aryan has *bā*-forms; but what the history of these central Dravidian and IA forms is, is still uncertain.

19. The same two lines of argument that were used in answering the first question in the affirmative will give an affirmative answer also to the second question, whether in PDr. kinship terms were preceded by pronominal possessive forms that were only plural. The *t*-forms with which we started are explicable only as relic forms of such a system. Moreover, Old Tamil, Kota, Kolami, Gondi, and Kuwi (especially as described by Schulze) have such a system for certain kinship morphemes; Kurukh, which has a peculiarly extensive system of inalienable possession, has only one morpheme (*bas* 'father') which requires a plural pronominal possessive.

Even though the system should be postulated for Proto-Dravidian on these two grounds, it is noteworthy that the living systems have few morphemes in common. Old Tamil, Kota, and Gondi, in fact, have no morphemes in their systems in common with one another or with the other three languages. Kolami and Kuwi seem to have a 'mother' morpheme in common (Kol. *am ay* 'my, our mother', Kuwi [S] *maiḥa*, *māḥa* 'my, our mother', *mīḥa* 'your mother'); already in Old Tamil *āy*, *tāy*, etc. are not recognizable as parts of the system seen in *empi*, *numpi*, *tampi*. Kolami, Kuwi, and Kurukh have a 'father' morpheme in common (Kol. *ba'n*, Kuwi *māba*, *mība*, Kur. *bas*); Tamil, Kota, and Gondi do not have this morpheme at all. It can be noted, however, that of the six languages all but Kota agree in having some morpheme for 'father' preceded by plural pronominal possessives (Old Tamil *-tai*, Gondi *āpō* or something like that, the other three *-bā*).

parison of this utterance with other utterances yields *ta-k* as the word meaning 'father', and any possessive pronominal form preceding it is a separate word. The resulting syntactic statements are also simpler than if there were no word division at this point.

20. The question of the meaning involved in the use of plural pronominal possessives must be discussed. I have been saying in previous paragraphs that the number of the possessor or possessors is not indicated. This is however to talk in terms of English translation. What is in fact given by the pronominal forms is a plural reference, and an attempt must be made to pin this reference down.

My first analysis of the Kolami data seemed to indicate that the possessed kinship terms all denoted kin older than, and consequently of superior status to, the EGO of the reference. The plural pronoun might in that case have been an honorific plural with reference to the possessed kin. However, it turned out on more searching analysis that *bai* 'sister' does not always have reference to a person of superior status and yet is preceded by the plural pronouns. Moreover, the Tamil forms quite certainly denote in two cases (*-pi* 'younger brother' and *-kai* 'younger sister') persons of inferior status.

The correct explanation is much simpler. The plural reference of the pronouns is not to either the possessed or the possessor individual, but to the whole family to which the kin denoted by the noun is related. Thus, Ta. *empi* is 'younger brother belonging to us, i.e. our (exclusive, i.e. not your) family', Kol. *am ba'n* is 'father belonging to us, i.e. our (exclusive, i.e. not your) family', Kur. *nimbis* is 'father belonging to you plural, i.e. your family'. The explanation is ethnologically convincing among communities in which the family, especially in the form of the extended family, is strongly structured. It points up especially the opposition between 'we exclusive' and 'you plural'. If the use of the reflexive plural is not so well motivated by the ethnological explanation, we can revert to linguistic analogy in the form of the oppositions *em- : tam-* and *nim- : tam-*.

21. Kota has a noteworthy construction with kinship terms that seems to be unparalleled elsewhere in the Dravidian family.

The demonstrative pronouns have each three forms denoting persons: *avn*, *ivn* 'that, this man', *avl*, *ivl* 'that, this woman', *avr*, *ivr* 'those, these persons' (Kota texts 1.24, §39). When, however, a demonstrative pronoun preceding a kinship term is a possessive attribute ('his father', 'her son', etc.), the only forms found are the masculine and feminine singular, and the demonstrative has the gender of the kinship term, i.e. the number and sex of the possessor or possessors are neutralized and are determinable only from the context. Examples: *avn ayn* 'his, her, their father', *avl av* 'his, her, their mother', *avl mo-l* 'his, her, their daughter', *avn kara-l* 'his, her, their younger brother', *avn angul* 'his, her, their elder brothers'. With some kinship terms (e.g. those denoting spouses and those used only of relatives or affines of a man or of relatives or affines of a woman) there can be no ambiguity about the sex of the possessor, even out of context. Examples: *avl ped* 'his wife', *avl pedgu-l* 'their wives', *avn a-l* 'her husband', *avn antamn* 'her, their brother-in-law (husband's brother or sister's husband)', *avn ayl* 'his, their brother-in-law (sister's husband or wife's brother)'.

22. Differential treatment of kinship terms which sets them off as a subclass of a larger class (usually the noun class) has been recorded for numerous languages. It is often of the first kind that we have identified for Proto-Dravidian, i.e. the subclass is made up of bound forms 'inalienably possessed'. Many Ameri-

can Indian languages have this feature, usually with names of body parts added to the kinship terms to form a unified subclass; examples are the Algonquian family, Chiricahua Apache and Chipewyan (perhaps all the Athapaskan languages), Eskimo, and Tunica.¹³ In Hopi and Aztec, both of the Uto-Aztecan family,¹⁴ the inalienably possessed subclass consists of kinship terms alone; in Tübatulabal, of the same family,¹⁵ it consists of kinship terms, a few names of body parts, and a few other nouns. In Chitimacha¹⁶ the special class consists of nouns referring to 'relationship' and to 'located parts'; the construction is a syntactic one. On the other hand, the differential treatment may consist of a special set of possessive affixes, different from those used by other noun classes. Thus, in Yuma¹⁷ kinship terms and names of body parts have a different set of possessive prefixes from those used by other nouns; in Takelma¹⁸ many kinship terms have a different set of possessive affixes from those of other noun classes. This small sample suggests that it might be profitable to map this trait for all American Indian languages for which we have data, to determine whether this is a trait to be added to those pointed out by Boas¹⁹ as diffused from language to language over a contiguous area, which need not be a small one.

The search for similar or other differential treatment of kinship terms in the Old World, in order to determine whether Proto-Dravidian is isolated in this respect or not, produces some interesting examples. In India, the Munda languages are badly described, and so far there is no clear report of anything like inalienable possession or of any other special treatment.²⁰ Burushaski in the far northwestern tip of the old Indian Empire provides a beautifully clear instance of special treatment of kinship terms, names of body parts, and certain other nouns, which form a subclass because of their inalienable possession, with possessive prefixes which are not otherwise used with nouns.²¹ It is reported that some Indo-Aryan dialects of the northwest of the Indian peninsula (e.g. Pashai) have a similar feature.²² Certainly the Indo-Aryan languages in general and the

¹³ The references for most of the American Indian languages mentioned are to *Linguistic structures of native America* (Viking Fund publications in anthropology 6; New York, 1946). Algonquian: Leonard Bloomfield, *LSNA* 96, §32; Chiricahua Apache: Harry Hoijer, *LSNA* 75; Chipewyan: Li Fang-kuei, *LSNA* 401, §12c; Eskimo: Morris Swadesh, *LSNA* 35; Tunica: Mary R. Haas, *LSNA* 358, §5:73; *Tunica* 64, §4.83 (New York, 1940).

¹⁴ Benjamin Lee Whorf, *LSNA* 171, §7.5; 381, §3:2.

¹⁵ C. F. Voegelin, *Tübatulabal grammar* 140 ff. (*UCPAAE* 34.2; Berkeley, 1935).

¹⁶ Morris Swadesh, *LSNA* 326, §3:42.

¹⁷ A. M. Halpern, *LSNA* 264, §23; *IJAL* 12.209, §27 (1946).

¹⁸ Edward Sapir, *The Takelma language of southwestern Oregon* 231-4 (*Handbook of American Indian languages*, Part 2; Bureau of American Ethnology, Bulletin 40; Washington, 1912).

¹⁹ Franz Boas, *Handbook of American Indian languages*, Part 1, 45 ff. (Bureau of American Ethnology, Bulletin 40; Washington, 1911); *Lg.* 5.1-7 (1929).

²⁰ P. O. Bodding, *A Santali grammar for beginners* 30-2 (Benagaria, India, 1929), records a set of pronominal possessive suffixes which do not distinguish the number of the possessor; they certainly occur with kinship terms and it is possible that they occur only with kinship terms, but Bodding is not explicit.

²¹ D. L. R. Lorimer, *The Burushaski language* 1.127-39, esp. §§128, 129 (Oslo, 1935).

²² Georg Morgenstierne, preface to Lorimer's work referred to in fn. 21, p. x. It does not seem that Morgenstierne has reported this elsewhere in more detail.

Iranian languages have no such feature, and these Indo-Aryan dialects are near enough geographically to Burushaski to allow us to posit borrowing from Burushaski at a period of contact. The great Sino-Tibetan family seems not to give these nouns a differential treatment, nor so far as I can find do the languages of Farther India. Of the Malayo-Polynesian family, the Melanesian and Micronesian sections²³ differentiate between a noun class consisting of kinship terms, names of body parts, and some names of parts of whole objects, and a second class consisting of all other nouns; the differentiation rests on different methods of expressing possession. In a few of the languages (e.g. Fiji and Saa) the nouns we are interested in are recorded as technically inalienably possessed, in that they do not occur without affixed pronominal possessives; this may be true of others of the languages also, but the evidence is not good enough to go very far with this feature. In some of the Australian languages, especially some of those in the Northern Kimberley district (on the far northwest coast),²⁴ nouns denoting parts of the body and some relationship terms have pronominal possessive prefixes, while other nouns take independent possessives; there are very unclear hints at inalienable possession in the published accounts. The differentiation in these languages reminds one of the Melanesian-Micronesian type of differentiation. Nothing of this sort, so far as my search goes, is found in the 'Papuan' languages.

We find, then, that two areas in India have inalienable possession. The area of Burushaski and northwestern Indo-Aryan is probably to be counted as historically one, with diffusion of this feature, probably from Burushaski to Indo-Aryan; the subclass of nouns here includes kinship terms and names of body parts. The Proto-Dravidian trait concerns only kinship terms. The Dravidian area is now widely separated from the area of Burushaski and northwestern Indo-Aryan (except for Brahui, which in any case does not have the Proto-Dravidian trait). Are we to posit two separate developments? Or is it preferable to think of diffusion in one direction or the other, at some early period when Proto-Dravidian speakers occupied territory much further north and were possibly in contact with Proto-Burushaski speakers? We are here in the realm of hypothesis, even of 'sheer' hypothesis, and I will not say more. It would be even more riskily hypothetical to think of an opportunity for diffusion that

²³ L. Lévy-Bruhl, *La possession dans les langues mélanésiennes*, *MSL* 19.96-104 (1914). Before I found this reference, my colleague C. Douglas Chrétien had called my attention to this Melanesian trait. The chief primary sources and discussions are to be found passim in: R. H. Codrington, *The Melanesian languages*, esp. 128 ff. (Oxford, 1885); Sidney H. Ray, *Reports of the Cambridge anthropological expedition to Torres Straits*, Vol. 3: *Linguistics*, esp. 437 (Cambridge, 1907); id., *A comparative study of the Melanesian island languages*, esp. 69 (Cambridge, 1926). For Micronesian I am indebted to my colleague Denzel R. Carr for an oral communication and the use of unpublished U. S. Navy textbooks.

²⁴ See e.g. A. Capell in *Studies in Australian linguistics* 59 (The Oceania Monographs, No. 3; Sydney, 1938), and A. Capell and A. P. Elkin, *ibid.* 99. I am indebted for this reference to Chrétien. The literature on the Australian languages is hard to control, and I must rely on Capell's seemingly exact statement (*loc.cit.*), the implication of which is that it is only these North Kimberley languages that have this feature. Certainly, Aranda as described by T. G. H. Strehlow (*Aranda phonetics and grammar*; The Oceania Monographs, No. 7; Sydney, n.d.; reprinted from *Oceania* 12-4, 1942-4) does not have it.

would embrace the Indian areas, Melanesia–Micronesia, and northern Australia. Certainly, independent invention cannot be ruled out for this linguistic trait. But, once this is posited, we are left with the insoluble question of how often it was invented. Perhaps more and better information from India and other parts of Asia will help towards at least a partial solution.

ETYMOLOGIES A

1. Ta. *akkā*, *akkai*, *akkan*, *akkāl*, *akkacci*, *akkaicci* *elder sister*; Ma. *akka id.* (rare); Kota *akn elder sister or female parallel cousin*; Toda *ok, okok id.*; Ka. *akka elder sister*; Koḍ. *akkē elder sister or female parallel cousin*; Tu. *akka, akkā elder sister*; Te. *akka id.*; Kol. (SR) *akkābāi id.*; Go. *tākkā*, (L) *akkā id.*

2. Ta. *aññai mother*; *annai mother, elder sister*; *emm-annai our mother*; *tamm-annai mother*; *tannai mother, elder sister*; Pa. *añña father's sister*.

3. Ta. *annan*, *annacci elder brother*; Ma. *annan id.*; Kota *aṇ elder brother or male parallel cousin*; Toda *oṇ, oṇoṇ id.*; Ka. *anna elder brother*; Koḍ. *annē elder brother or male parallel cousin*; *tammaṇē younger brother*; Tu. *anne elder brother, maternal uncle, an elderly man*; Te. *anna elder brother*; Kol. (Haig²⁵) *annāk id.*; Go. *tānnāl id.*; *tānhārāl id.*, *respectful*.

4. Ta. *avvai mother, old woman, woman ascetic*; Kota *av mother, mother's sister or female parallel cousin*; Toda *af, afuf id.*; Ka. *avve, avva mother, grandmother, any elderly woman*; Koḍ. *avvē mother, mother's sister or female parallel cousin*; Te. *avva mother, grandmother, an old woman*; Go. *āwṃāl, āwhārī mother, dam of man or beast*; *maiṃāl my mother*; *mīṃāl thy mother*.

5 (a). Ta. *āy, āyi, yāy, ñāy mother*; *āycci id., grandmother*; *āyāl id., old woman*; Ka. *āyi mother*; Kol. *ay (? a·y) id.*; Oll. *aya id.*; *ayal woman, wife*; Go. *yāyāl, (M) yāyo mother*; Kui *aia, aiali id., woman*; Kuwi (F) *iya mother, maiya my mother, miya your mother, aiya woman*; (S) *iya (j = y) mother, woman, māiya my, our mother, mīya your mother*; Kur. *ayang mother* (without explicit reference to the children); *ayō mother* (*ingyō, ningyō, tangyō my, thy, [his, her, their] mother*; *yō vocative*); Malt. *ayya my mother*. Cf. Assamese *āi mother*, Bengali *āi mother's mother or aunt*, Oriya *āi mother's mother*, Sindhi *āi mother, aunt*, Gujarati *āi mother, grandmother*, Marathi *āi mother*; see Turner s.v. *āimāi*.

5 (b). Ta. *tāy, tāyār mother*; *tāymai motherhood, motherliness*; Ma. *tāyi mother*; Toda *to·y mother* (in poetry); Ka. *tāy, tāyi, tāye mother*; Koḍ. *ta·yi grandmother*; Tu. *tāyi mother*; Te. *tāyi id.*

6. Ta. *taṅkai, taṅkaicci, taṅkacci younger sister or female parallel cousin*; *eṅkai, eṅkaicci my younger sister*; *nuṅkai, uṅkai your younger sister*; *taṅkāi younger sister*; *kai, kaiyai id.* (lexical); Ma. *taṅga, taṅgacci younger sister*; Ka. *taṅgi id.*; Koḍ. *tange younger sister or female parallel cousin*; Tu. *taṅgi, taṅgaḍi male's younger sister*; Kui *angi, tangi younger sister*; Kuwi (F) *tangi sister*.

7. Ta. *tantai father*; *entai, nantai (Caldwell) my father, our father, my elder brother*; *nuntai, untai your father*; Ma. *tanta father, also of animals*; Ka. *tande father*; Te. *taṇḍri id.*; Pa. *tend id.*; Kui *tanji id.*; Kuwi (S) *tanzi (z = j) id.*

²⁵ Wolseley Haig, A comparative vocabulary of the Gōṇḍī and Kōlāmī languages, *Journal of the Asiatic Society of Bengal* 66.1.185–91 (1897).

Ta. Ma. Ka. agree on the PDr. dental combination *n-t- required by our hypothesis of a morpheme (-)tai. Te. and Kui-Kuwi can only be unified on the basis of PDr. *nr, i.e. the corresponding alveolar cluster. Pa. nd could apparently correspond to either. Is this a case of two PDr. dialectic forms?

8. Ta. *tampi* younger brother or male parallel cousin; *empi* my younger brother; *numpi*, *umpi* your younger brother; Ma. *tampi*, *tampān* younger brother; Ka. *tamma* id.; Koḍ. *tammanē* (voc. *tamayya*ⁿ) id.; Tu. *tammaiya* an affectionate form of addressing a younger brother; Te. *tammūḍu* younger brother; Kol. (SR) *iltāmā* younger sister's husband; Go. *tāmmūr* younger brother; Kui *ambesa*, *tambesa* id. Probably the forms with *tamp-/tamb-* and those with *tamm-* do not belong together closely, if at all.

9. Ta. *tallai* mother; Ma. *talla* id.; Te. *talli*, *tali* id.; Pa. *tal* id.; *tal vanda* thumb; Kui *ṭaḍi* mother, woman, female; *tali* female bird or animal, hen; Kuwi (S) *talli* mother; *talli vānsu* thumb; (F) *tāli vānjū* thumb.

10. Kol. *ba'n* (so-and-so's) father (voc. *ba'*); *doo ba'k* grandfather; Kur. *bas* (so-and-so's) father (voc. *bā*). Cf. Marathi *bā* father! and the equivalent in a few other IA languages. Cf. Te. *abba* father; Kui *ābā* father, grandfather, ancestor, an elder, a superior; Kuwi (F) *abba* father, *māba* my father, *miba* your father; (S) *āba* father, *māba* my, our father, *miba* your father; Kur. *abbā* father; Malt. *abba* id.; Br. *abā*, *abbā* father, grandfather, respectful term of address (cf. Balochi *abbā*, Sindhi *abō*); all these are somehow connected with IA material.

11. Ta. *aṭu* (*aṭuv-*, *aṭṭ-*) smite, kill, destroy, conquer; *āṭu* n. killing, ruining, victory; *aṭi* (*aṭipp-*, *aṭitt-*) beat, strike, dash against, defeat, overcome, slay, punish; n. blow, stroke; Ma. *aṭi* n. blow, stroke; *aṭikka* beat, strike; Kota *ayr-* (*aṛ-*) strike (ball in game); Ka. *aḍacu* rap, cuff; *aḍi* beating, a blow; Koḍ. *aḍi-* (*aḍip-*, *aḍic-*) (lightning) strikes; Te. *aḍucu* beat, strike, break, destroy; Pa. *aṭṭ-* strike; Oll. *aṭ-strike*, (rain) comes in torrents.

12. Ta. *taku* (*takuv-*, *takk-/takunt-*) be fit, suitable, proper, worthy, adequate, be excellent; *takavu* suitability, fitness, worthiness, eminence, morality; *takal* fitness; *takalōn* one who is eminently fitted, suitable, or worthy; *takutai*, *takai* fitness, meetness, suitability, adequacy, worthiness, excellence, beauty; *takai* (*takaiv-*, *takaint-*) be beautiful; *takaimai* fitness, excellence, beauty; *takku* (*takki-*) be fit, suitable, becoming; *takkatu* that which is fit or proper; *takkār*, *takkōr* worthy persons; *takappan* father; Ma. *taku* be fit, suit; *takka* adj. fit; *takkavan* suitable, proper; *takkam* fitness, convenient time; *takappan* father; Toda *tokoθ* suitable; Ka. *tagu* be fit or proper, suit; *takku* fit, proper, right, deserving, suitable, good; a good man, a friend; *takkame* fitness, etc.; Tu. *takka* fit, suitable, proper, deserving, worthy; Te. *tagu* be proper, becoming, fit, suitable, be worthy, deserving, competent; *tagina* proper, becoming, fit, suitable.

ETYMOLOGIES B

Evidence for Kota simplification of homorganic nasal plus stop by loss of nasal.

*mp > Kota b.

Ta. *erumpu*, *uravi* ant; Ma. *erumpu*, *urumpu*; Kota *irb*; Toda *irb* (r instead of r!); Ka. *irumpu*, *irupe*; Koḍ. *urupī*.

Ta.-Ma. kompu *branch, horn of animal, musical horn, tusk*; Kota kob *id.*; Toda kub *horn blown by Kota musician*; Ka. kombu *branch, horn of animal, musical horn, tusk*; Koḍ. kombī *branch, horn of animal*; Tu. kombu *id., musical horn, tusk*; Te. kommu *id.*; komma *branch*; Kol. kom *branch, horn*; Nk. komm *horn*; Kuwi (S) kommu *id., komma branch*.

Ta. tummu (tummi-), tumpu (tumpi-) *to sneeze*; Ma. tummuka, tumpuka *id.*; Kota tub- (tuby-) *id.*; Toda түb- (tüby-) *id.*; Koḍ. tımm- (tımmi-) *id.*; Tu. tumbılı *a sneeze*; Te. tummu *to sneeze; a sneeze*; Kol. tum- (tumt-) *to sneeze*; Nk. tum *a sneeze*; Pa. tumm- *to sneeze*; Oll. tum- *id.*; Kuwi (F) tühmali, (S) tūminai *id.*; Kur. tum'nā *id.*; Malt. tume *id.*

Kota tub- (tuby-) *be filled full*; Ka. tumbu *become full, filled up*; Koḍ. dumb- (dumbi-) *become full*; Tu. tumbuni *be filled*. This is a particularly valuable example, since Toda does not have the word and Kota is therefore independent of Toda; in some of the other examples there may be borrowing between the two languages in one direction or the other (cf. the words for 'horn'; Toda has borrowed from Kota).

Ta. narampu *nerve, tendon, sinew, blood-vessel*; Ma. ñarampu, narampu *id.*; Kota narb *muscle, sinew*; Toda narb *muscle, vein*; Ka. nara, naravu *sinew, tendon, nerve, vein, artery*; Tu. nara *vein, nerve, tendon*; narambu *sinew, nerve, pulse*; Te. naramu *vein, artery, nerve, tendon*; Kol. (Kin.) naram *vein*; Pa. nerub (pl. nerbul) *vein*; Kui ḍrāmbu *tendon, sinew*; Kuwi (S) naromi *nerve*; Kur. narī *pulse* (? < IA); Malt. nāru *the veins*.

Ta.-Ma. pāmbu *snake*; Kota pa'b; Toda po'b; Ka. pāvu; Koḍ. pa'mbī; Tu. hāvu; Te. pāmu; Kol.-Nk. pa'm; Pa.-Oll. bām. Cf. Prakrit (Deśināmamālā) pāva- *snake*.

*nt > Kota d. Many examples of past tense stems which in Ta. have -nt-; e.g. eṛu (eṛuv-, eṛunt-) *rise, ascend, fly*; Kota eṛv- (eṛd-) (*smoke*) *raises up, spring up from position, fly*; Toda öḍ- (öḍθ-) *be high or raised, rise, get up*; Ka. eṛ (eṛd-), eṛu (edd-) *stand up, rise*; Koḍ. ē- (ēv-, ēdd-) *get up*.

*nr > Kota ḍ; *nri > Kota j.

Ta. en (enp-, enr-) *say (so-and-so)*; Ma. ennuka; Kota in- (iḍ-); Toda in- (iḍ-); Ka. en- (end-), an- (and-); Koḍ. *enn- (emb-, end-); Tu. anpini, inpini; Te. anu; Kol. en- (ent-); Pa. en- (ett-, end-); Go. Indānā (Inj-); Kui in-b-a (is-); Kuwi (F) injali, (S) innai; Kur. ānnā; Malt. áne.

Ta. tin- (tinp-, tinr-) *eat*; Ma. tinnuka; Kota tin- (tiḍ-); Toda tīn- (tiḍ-); Ka. tin (tind-); Koḍ. tinn- (timb-, tind-); Tu. tinpini; Te. tinu; Kol.-Nk.-Pa. tin (tind-); Oll. tin- (tiy-, tiḍ-); Go. tīndānā (tīnj-); Kui tin-b-a (tis-); Kuwi (F) tinjali, (S) tinnai; Kur. tindnā *put into another's mouth*; Malt. tinde *feed by hand*.

Ta. panri *pig*; Ma. panni; Kota paj; Toda pody (should have ḍ; probably borrowed from a modern Tamil dialect with -nt-); Ka.-Koḍ. pandi; Tu. pañji; Te. pandi; Pa. pend, (NE) penḍ; Oll. paṇḍ; Go. pāddī; Kui paji; Kuwi (F) pajji, (S) pazzi (z = j).

*ṇt > Kota ḍ.

Ta. āl (ālv-, āṇṭ-) *rule, reign over, control or manage (as a household), keep or maintain in use*; Ma. āluka *rule, possess, have*; Kota a-ṇ- (a-ḍ-) *possess, rule, own*,

keep; Toda o·l- (o·ḍ-) *own (buffaloes), rule*; Ka. āl (ālḍ-) *get, possess, rule, manage*; Koḍ. a·l- (a·li-) *rule*; Tu. āluni *rule*; Te. ēlu *rule, control, manage*.

Ta. uṇ (uṇṇ-/uṇṇuv-, uṇṭ-) *eat or drink, suck*; Ma. uṇṇuka *eat (esp. rice), suck*; Kota uṇ- (uḍ-) *drink, suck*; Toda uṇ- (uḍ-) *id.*; Ka. uṇ (uṇḍ-) *eat a meal, suck*; Koḍ. uṇṇ- (umb-, uṇḍ-) *eat a meal*; Tu. uṇupini, uṇṇini *id.*; Kol. un- (und-) *drink*; Nk.-Pa. un- (uṇḍ-) *id.*; Oll. un- *id.*; Go. ūṇḍānā *id.*; jāwā ūṇḍānā *take food*; Kui uṇ-b-a (uṭ-) *drink, partake of (a meal)*; Kuwi (F) ūṇḍali *drink*, (S) unnai *drink, suck*; Kur. ōṇnā (oṇḍas) *drink, eat (rice)*; Malt. ōne (oṇḍ-) *drink*.

*ṇk > Kota g.

Ta. amuṇku (amuṇki-) *sink, be pressed down, crushed*; Ma. amuṇṇuka *id.*; Kota amg- (amgy-) *be pressed hard*; Toda omx- (omxy-) *be pressed down*; Ka. amugu *yield to pressure*.

Ta. kuraṇku *monkey*; Ma. kuraṇṇu; Kota korg; Toda kwarg; Ka. koraṇgi; Tu. kuraṇga.

Ta. tiṇkaḷ *moon, month*; Ma. tiṇkaḷ *moon*; tiṇṇaḷ *month*; Kota tigḷ *moon*; tigḷ *month* (probably a borrowing); Toda tigīḷ *moon*; ti·l *month*; Ka. tiṇgaḷ *moon, month*; Koḍ. tiṇga *month*; Tu. tiṇgoḷu *moon, month*; Kui tingal danju *crescent moon*.

A BURMESE JĀTAKA COMMENTARY

WILLIAM S. CORNYN

Yale University

Among the Burmese versions of the Pāli Jātaka there is one, the *nga:ya. nga:ze* 'the five hundred and fifty',¹ which is a faithful rendition of the original and which retains the gathas in a Burmanized Pāli. These gathas are accompanied by Burmese commentaries, which present points of linguistic interest both in the Pāli and in the Burmese.

The gathas cited here are taken from the Sasa-Jātaka,² called in Burmese the *tha.tha. pan-di.ta. za*.³ The first of these is, in the Pāli:

*Satta me rohitā macchā udakā thalam ubbhata,
idaṃ brāhmaṇa me atthi, etaṃ bhutvā vane vasā.*

In its Burmanized form this gatha is recited as follows:

*tha'ta. mei ro:hi.ta myi'hsa, qu.daka htala. mou'bata,
qi.dan byamana. mei qa'hti., qei:tan bou'twa wanei watha.*

The commentary follows immediately in a word-for-word translation, the order of the glossed forms being so arranged that a meaningful translation emerges. It runs as follows: *byamana*.—*poun-na*: 'Brahmin', *mei—nga.qa*: 'to me', *qu.daka—yei-hma*. 'from water', *htalan—ci:goun:dhon*. 'to land', *qou'bata—tanga-dhi hse tin qa' koun-dho*: 'which have been brought out by a fisherman', *tha'ta.—hkun-nakaun koun-dho*: 'seven', *ro:hi.ta. myi'hsa—ngajin:dou.dhi* 'large gudgeon (pl.)', *than-ti.—shi. koun-qi*. 'there are'. *qi.dan—qi qou'sa-dhi* 'this thing', *mei—nga.qa*: 'to me', *qa'hti.—shi.qi*. 'there is'. *qei:tan—qi qou'sa-gou* 'this thing', *bou'twa—sa:ywei*. 'eating', *wanei—to:hnaï* 'in the forest', *watha.—nei-lo*. 'remain'.

Taking only the Burmese glosses, the commentary runs: *poun-na*: *nga.qa*: *yei-hma. ci:goun:dhon. tanga-dhi hse tin qa' koun-dho: hkun-nakaun koun-dho: ngajin:dou.dhi shi. koun-qi./ qi qou'sa-dhi nga.qa: shi.qi./qi qou'sa-gou sa:ywei*.

¹ The edition of *nga:ya. nga:ze* from which these passages are quoted is that of the Zabu Meitswe Press, 5 volumes (Rangoon, 1908-12); they are from 2.612-7.

The transcription used in this paper differs in the following points from that of the author's *Outline of Burmese grammar* (Lang. Diss. No. 38, 1944): the aspirates are written with the *h* first: *hp, ht, hk, hc, hs* for *ph, th, kh, ch, sh*; the digraphs *th, dh, sh* are written for the spirants, and *ng* for the velar nasal; *q* is written for an initial glottal stop; the tone marks are written with marks of punctuation: tone I hyphen, II colon, III period, IV apostrophe, atonic syllables unmarked. Initial sandhi replacements are shown, but the assimilation of finals is not. The hyphen for tone I is omitted when a space follows. These changes are made only to simplify the writing; they do not imply a change in the phonemic analysis.

² Jātaka 316, quoted from V. Fausbøll (ed.), *The Jātaka together with its commentary* 3.51-6 (London, 1883). Also in Dines Andersen, *A Pāli reader with notes and glossary* 14-6 (London, 1901).

³ For the recitation of the gathas I am indebted to my friend U Wun of the University of Rangoon. They were recorded in connection with a reading of the tale from the collection of U Ponnya, *Wuttu Baung Gyoke* 62-73 (Rangoon, 1950).

to:hnaï' nei-lo./ 'Oh Brahmin, I have seven gudgeon which a fisherman has brought out from the water onto land. This I have. Eat this and remain in the forest.'

The translation of the first *mei* (Pāli *me*) by *nga.qa:* 'to me', i.e. as a Pāli dative instead of the usual interpretation as a Pāli instrumental, gives rise to certain difficulties which the commentary resolves by adding *than-ti*. (Pāli *santi*), which does not occur in the gatha. In addition, the commentary renders *gou'bata* (Pāli *ubbhatā*) as *tanga-dhi hse tin qa' koun-dho:*, in which the notion of *tanga-dhi* 'fisherman' has been added to what would have been an accurate translation: *hse tin* 'take out and put upon' plus *qa'* as a sign for the Pāli passive, *koun* as the sign of the Pāli plural, and *-tho:* as the sign for the Pāli adjective, marking the whole as a plural passive participle. It is possible that the commentary introduces the fisherman and makes the other changes in order to avoid ascribing to the otter, the future Ānanda, the crime of fishing. Another point of interest in this passage is that the *htala. mou'bata* of the gatha, clearly so marked and recited, is in the commentary *htalan* and *gou'bata*, corresponding to the Pāli *thalam ubbhatā*.

The second of the gathas reads, in Pāli:

*Dussa me khetpālāssa rattibhattam apābhataṃ,
māmasulā ca dve godhā ekan ca dadhivātrakam,
idaṃ brāhmaṇa me atthi, etaṃ bhutvā vane vasā*

In its Burmanized form this gatha is recited as follows:

*dou'tha. mei hki'ta. pa-la'tha., ra'tein-ba'tan qa.pa-batan,
man-tha. thu-la-sa. dwei go:dan qei:kin-sa. da.di.wa-rakan
qi.dan byamana. mei qa'hti., qei:tan bou'twa wanei watha.*

The commentary runs as follows: *byamana.—poun-na:* 'Brahmin', *dou'tha.—qi gamyi shi.dho:* 'having this name', *hki'ta.pa-la'tha.—le-zaun.qi.* 'of the field-watcher', *ra'tein-ba'tan—nyi.za-gou* 'evening meal', *mei—nga-dhi* 'I, by me', *qa.pa-batan—hsaun-ge.qa.qi.* 'has been brought', *dwei—hnaḥku. doun-dho:* 'two', *man-tha. thu-la-sa.—game: tazou.dou. lagaun:* 'spits of meat and', *go:din-sa.—hpu'kin lagaun:* 'a lizard and', *qei:kan—tahku.dho:* 'one', *da.di.wa-rakinsa.—nou.dan:gou: lagaun:* 'a jar of milk curds and', *qi.dan—qi gou'sa-dhi* 'this thing', *mei—nga.qa:* 'to me', *qa'hti.—shi.qi.* 'there is', *qei:tan—htou gou'sa-gou* 'that thing', *bou'twa—sa.ywei.* 'eating', *wanei—to:hnaï* 'in the forest', *watha.—nei-lo.* 'remain'.

Taking only the Burmese glosses, the commentary on the second gatha runs: *poun-na:, qi gamyi shi.dho: le-zaun.qi. nyi.za-gou nga-dhi hsaun-ge.qa'qi./ hnaḥku. koun-dho: game: tazou.dou. lagaun:, hpu'kin lagaun:, tahku.dho: nou.dan:quo: lagaun:/ qi gou'sa-dhi nga.qa: shi.qi./ htu gou'sa-gou sa.ywei., to:hnaï' nei-lo./* 'Oh Brahmin, the night meal of this field-watcher has been brought by me. Two spits of meat, a lizard, and a jar of milk curds. This I have. Eat that and remain in the forest.'

dou'tha. (Pāli *dussa*) is glossed as 'having this name', i.e. an indication of a demonstrative pronoun. Andersen⁴ notes this interpretation as present in the

⁴ Andersen, *A Pāli reader* 124.

Singhalese mss., but prefers the interpretation *dussa* = Skt. *dashya* 'corruptible, easily to be spoiled or damaged'. He rejects the translations 'wrongfully' in Francis and Neil⁵ and 'boshast' in Dutoit.⁶ Another interpretation is that *dou'tha* is the name of the field-watcher. This is unlikely, since the normal Burmese for such an interpretation would be *qi dou'tha. qamyi shi.dho*: 'having the name *dou'tha*'. The gloss of *mei* (Pāli *me*) in the first line of this gatha by *nga-dhi* is of interest when compared with the similar form in the first gatha, where it is glossed by *nga.qa*:. The particle *-thi* in this case marks the source of the action and indicates the Pāli instrumental. The commentary does not agree with the gatha in its reporting of the forms with *-sa*. (Pāli *ca*). The gatha has *man-tha. thu-la-sa.* and *gei:kin-sa.*; the commentary has *man-tha. thu-la-sa., go:diin-sa.,* and *da. di. wa-rakin-sa.* The last lines of the first and the second gathas are identical, but the gloss of *gei:tan* in the second is *htou gou'sa-gou* 'that thing', and in the first *qi gou'sa-gou* 'this thing'.

The third of the gathas reads, in Pāli:

*Ambapakk' odakam sīlam sītaccāyam manoramam,
idaṃ brāhmaṇa me atthi, etaṃ bhutvā vane vasā*

In its Burmanized form this gatha is recited as follows:

*gan-ba.pe'ko:dakan thi-tan, thi-ti'hsa-yan mano: raman,
qi.dan byamana. mei qa'hti., gei-tan bou'twa wanei watha.*

The commentary runs: *byamana*.—*poun-na*: 'Brahmin', *mei—nga.qa*: 'to me', *gan-ba.pe'kan—thaye'thi:hme. lagaun*: 'ripe mangoes and', *thi-tan—gei:dho*: 'which is cool', *qu.dakan—gin.ga yei lagaun*: 'Ganges water and', *mano: raman—hnaloun: mwei. lyo-bwe shi.dho*: 'which is pleasing to the senses', *thi-ti'hsa-yan—hcan: gei:dho: qayei' lagaun*: 'shade which is cool and', *qi.dan—qi hsou-qa' pyi:dho: gou'sa-dhi* 'this thing which has been mentioned', *mei—nga.qa*: 'to me', *qa'hti.—shi.qi*. 'there is', *gei:tan—thaye'thi:hme. gin.ga yei-gou* 'ripe mangoes and Ganges water', *bou'twa—thau'ywei.* 'eating and drinking', *wanei—hnaloun: mwei. lyo-bwe qayei' shi.dho: to:hna'* 'in the forest which has shade pleasing to the senses', *watha.—nei-lo.* 'remain'.

Taking only the Burmese glosses: *poun-na*:, *nga.qa: thaye'thi:hme. lagaun*:, *gei:dho: gin.ga yei lagaun*:, *hnaloun: mwei. lyo-bwe shi.dho: hcan: gei:dho: qayei' lagaun*:, *qi hsou-qa' pyi:dho: gou'sa-dhi nga.qa: shi.qi./ thaye'thi:hme. gin.ga yei-gou sa: thau'ywei.*, *hnaloun: mwei. lyo-bwe qayei' shi.dho: to:hna'* *nei-lo./* 'Oh Brahmin, I have ripe mangoes and cool water of the Ganges and cool shade which is pleasing to the senses. This which has been mentioned I have. Eating and drinking the ripe mangoes and the water of the Ganges, remain in the forest which has shade pleasing to the senses.'

The first form of this gatha is recited either as quoted above, *gan-ba.pe'ko: dakan* agreeing with the Pāli, or as *gan-ba.pe'kan dakan*-. The phrase *mano: raman* is also heard *mano: rama*-. The first occurrence of *mei* in the commentary is unjustified either by the gatha or by the sense of the commentary. It is perhaps an

⁵ H. T. Francis and R. A. Neil (transl.), *The Jātaka or stories of the Buddha's former births* 3.34-7 (Cambridge, 1897).

⁶ Julius Dutoit, *Jātakam* 3.63 (München, 1911).

echo of the pattern of the first two commentaries. There are certain liberties taken in this commentary which are not paralleled in the others. *qan-ba.pe'kan* is glossed as 'RIPE mangoes,' *qu.dakan* is here 'GANGES water' instead of the plain 'water' of the first commentary. The second line of the gatha is identical with lines in the first two gathas but is expanded in the commentary so that this is no longer a simple gloss of the forms but rather a poetic restatement of the sense. *qi.dan*, in the first two commentaries 'this thing', is now 'this thing which has been mentioned.' *qei:tan*, which was also 'this thing', is now 'ripe mangoes and Ganges water'. *bou'twa* was 'eating' and is now 'eating and drinking'; *wanei* was 'in the forest' and is now 'in the forest which has shade pleasing to the senses'. This development foreshadows a later type of commentary, such as that of U Ponnya,⁷ where the word-for-word gloss is ignored and the gatha is recast and expanded in Burmese. It is interesting to note that the commentary of the fourth gatha returns to the style of the first two.

The fourth of the gathas reads, in Pāli:

*Na sasassa tilā atthi na mugga napi taṇḍula,
iminā agginā pakkamā mamaṃ bhutvā vane vasaṃ*

In its Burmanized form it is recited as follows:

*na. tha.tha'tha. ti.la qa'hti., na. mou'ka na-pi. tan-du.la,
qi.mi.na qe'ki.na pe'kan, man-than bou'twa wanei watha.*

The commentary runs as follows: *byamana.*—*poun-na:* 'Brahmin', *tha.tha'tha.*—*youn-qa:* 'to a rabbit', *ti.la—hnan:dou.dhi* 'sesamum seed', *na.qa'hti.—mashi.goun* 'there is not', *mou'ka—pe:nau'tou.dhi* 'beans', *ha.qa'hti.—mashi.goun* 'there is not', *tan-du.la—hsan-dou.shi* 'rice grains', *na. qa'hti.—mashi.goun* 'there is not', *qi.mi.na qe'ki.na—qi mi: hypin.* 'by means of this fire', *pe'kan—ce'tho:* 'which is cooked', *man-than—qatha:gou* 'flesh', *bou'twa—sa:ywei.* 'eating', *wanei—to:hnaï* 'in the forest', *watha.—nei-lo.* 'remain'.

Taking only the Burmese glosses: *poun-na:*, *youn-qa:* *hnan-dou.dhi mashi.goun*, *pe:nau'tou.dhi mashi.goun*, *hsan-dou.shi mashi.goun/ qi mi: hpyin. ce'tho: qatha:gou sa:ywei., to:hnaï nei-lo./* 'Oh Brahmin, a rabbit has no sesamum, no beans, no rice. Eat the flesh cooked by this fire and remain in the forest.'

The *byamana.* of this commentary is not represented in the gatha but parallels the use in the earlier ones indicating the person addressed. The *tha.tha'tha.*, genitive in the Pāli, is glossed *youn-qa:*, with the particle *-qa:* which regularly describes a Pāli dative, cf. the earlier *nga.qa:*. The Pāli genitive is regularly shown by the particle *-qi.*, as in the *le-zaun.qi.* of the second commentary. The Burmanized gatha has *man-than* (Pāli *mamsam*) 'flesh' for the *mamaṃ* 'me' of the Pāli. The gloss of *na. qa'hti.* by *mashi.goun* is difficult to interpret. The flavor of the Burmese form is archaic; at first glance it is to be taken as specifically plural, with the plural forms *hnan:dou.*, *pe:nau'tou.*, and *hsan-dou.*, cf. *shi. koun-qi.* for *than-ti* in the first commentary. But the fact that *mashi.goun* glosses *na.qa'hti.* leads to the conclusion that it is simply emphatic: 'there is not at all'.

This brief exposition of one Jātaka commentary shows the possibilities of such

⁷ Cf. fn. 3.

a line of investigation. There is first the elucidation of the gathas themselves: often a fresh interpretation, even where not acceptable in itself, will provide an insight into their meaning. There is further the information to be gained from the attempt to reflect Pāli syntactic features in a language of different structure; this is of course of greater interest to the student of Burmese than to the student of Pāli. Finally there is the light shed throughout these commentaries on cultural attitudes. Much can be gained for an understanding of the Burman and his view of life by studying how he reports the events of these tales.

DEMPWOLFF'S *R

ISIDORE DYEN

Yale University

1. Proto-Malayo-Polynesian *R was reconstructed by Dempwolff¹ to account for the correspondence of Tagalog *g* with Toba-Batak *r* and the lack of a corresponding phoneme in Javanese (i.e. loss of a consonant and contraction of the abutting vowels); thus, **beRas*, Tg. *bigás*,² Jv. *wòs*, TBt. *boras* 'polished rice'. Certain contradictions to Dempwolff's formula appeared not only in the Javanese cognates, but in those of Ngaju-Dayak and Merina,³ a dialect of Malagasy. The present article will examine these contradictions, to see whether they suggest hitherto unrecognized PMP distinctions.

2. Dempwolff's method of attack on the MP languages was first to work out inductively a system of proto-phonemes for the three languages listed above. His second step was to apply that system deductively to a series of other MP languages: Malay, Ngaju-Dayak, Merina, Fiji, Saa, Tonga, Futuna, Samoa. It is clear that any error in the original formulation would result, in the deductive application, in the appearance of systematically corresponding but inexplicably divergent reflexes. It will be shown in this article that Dempwolff's procedure was essentially insufficient (not 'wrong'). A third consolidating step is necessary which in principle is not different from the first step. In this third step, the reflexes of the languages newly introduced into comparison are treated on a par with the reflexes of the languages originally introduced. Dempwolff's limitation of his work to three languages turns out then to be a matter of convenience, and the results obtained can then be viewed as accurate within the limits of the procedure.

3. It is not only the errors in the original formulation which result in the appearance of inexplicably divergent reflexes. The effects of analogic change and of borrowing may be such as to produce like results. At a given stage in the reconstruction of a proto-language, the researcher may not be able to separate clearly the effects of secondary changes (analogic change and borrowing) from

¹ In his *Vergleichende Lautlehre des austronesischen Wortschatzes*, *Zeitschr. f. Eing.-Spr.*, Beihefte 15 (1934), 17 (1937), 19 (1938). (These will be referred to as Dempwolff 1, 2, and 3 respectively.) For a fuller reference and an account of the equivalences between our transcriptions, see Dyen, *The Malayo-Polynesian word for 'two'*, *Lg.* 23.50 fn. 1 (1947). For other differences, see Dyen, *Proto-Malayo-Polynesian *Z*, *Lg.* 27.534-40, esp. fn. 3 (1951).

² Tg. *biggás* 'polished rice' beside *bigás* is probably a product of contamination with *diggás* 'third polishing (of rice)'; note Hiligaynon Bisayan *bugás* 'hulled rice', *diggás* 'pound, grind well'. The contamination may have been favored first by the near homonymy of *bigás* and *diggás* resulting from the change of **e* to Tg. *i* (but to HlBs. *u*), and second by the loss of the postconsonantal *q* in some Tagalog dialects, which resulted in *digás*; thus *biggás* could have arisen at the *diggás*-*digás* isogloss.

³ I now use Merina instead of, though equivalent to, Dempwolff's Hova, thus following the practice of French scholars and of Otto Chr. Dahl, *Malgache et maanjan* (Oslo, 1951). For a few words on this point, see his p. 8.

primary change (phonemic change). A third purpose of this article is to point out that under such circumstances the researcher can advantageously have recourse to tentative or 'problematic' reconstructions.

4. Reconstructed PMP morphemes show no more than two consonants between vowels. Since no consonant clusters appear in initial or final position, it seems easy to reconstruct a syllable shape (C)V(C). The favorite shape of a root-morpheme was dissyllabic, i.e. (C)V(C)(C)V(C).

5. Since, furthermore, alternations of the initial consonant of a root are common in many of the Malayo-Polynesian languages so far systematically compared, it can be supposed that the initial consonant was frequently subject to analogic change. Since a medial or final consonant of a root was usually not a member of such alternations, consonants in these positions are in general not subject to analogic change. If however a language has undergone a change affecting its final consonants, an alternation may arise thereby which depends on the presence or absence of a suffix, and such consonants then too become subject to analogic change. Such considerations make intervocalic position within a root-morpheme (i.e. -VCV-) of paramount importance in the discovery of the consonant distinctions of PMP.

6. Besides analogy, the other change which makes for difficulties in the reconstruction of proto-phonemes is borrowing. When many words have been borrowed, the loanwords may systematically correspond with their etyma in much the same way that cognates correspond. If the borrowing occurs between related languages, then words related by borrowing may give the appearance of corresponding systematically, but at the same time at least one of their correspondences may differ from formulas established by other comparisons.⁴

7. On the other hand a proto-phoneme can yield phonetically similar results in related languages. If phonetically similar phonemes occur in the words of related languages in contact, and their correspondence does not agree with any formula that has been otherwise established, it is not safe to conclude that the words are related only by borrowing. If such a correspondence is temporarily attributed to the effects of borrowing, it may be discovered later that this explanation is insufficient. The situation in a given case, however, may be so complicated that it is not desirable to proceed immediately to the reconstruction of proto-phonemes.

8. There are four intervocalic correspondences which I have found subsumed under Dempwolff's *R. In each of the four, the Tagalog member is *g* and the Malay member is *r*. The essential members of each correspondence are given below, with the symbol X meaning 'loss':

*R ₁	Jv. X	NgD <i>h</i>	Mlg. ⁵ X
*R ₂	Jv. X	NgD <i>h</i>	Mlg. <i>z</i>
*R ₃	Jv. <i>r</i>	NgD <i>h</i>	Mlg. <i>z</i>
*R ₄	Jv. <i>r</i>	NgD <i>r</i>	Mlg. <i>r</i>

⁴ Otherwise the distinction between borrowed words and inherited words may not be discoverable without resort to earlier records, since morph-alternation in most of the MP languages so far systematically compared works in such a way that loanwords can easily acquire the alternations of inherited words.

⁵ For Malagasy. From Dahl's statement (*Malgache et maanja* 6) that the individuals

9. Not all of these correspondences are of equal value. The first is the soundest, even though Dempwolff treats the Merina reflex as exceptional from his *R.⁶ The second differs from the first only in the Merina reflex (here *z*). There is no presently known difference in the positions occupied by the two reflexes; but the instances are such that the Merina reflexes may yet turn out to be positionally determined variants. For this reason it seems desirable neither to proceed immediately to the reconstruction of different proto-phonemes nor on the other hand to ignore the difference in the Merina reflexes. The difference is therefore treated as 'problematic': the first is reconstructed as *R₁ and the second as *R₂.

10. The third correspondence (*R₃) differs from the second only in the Javanese reflex, here *r*. Since the Malay reflex is also *r*, and since Javanese and Malay have been in long-standing contact, the Javanese words must be suspected of being borrowings. It seems to me, however, that no clear-cut decision can be reached.

11. The fourth correspondence (*R₄) differs from each of the other three in at least two members (NgD, Mer. *r*) and therefore seems to be almost certainly separate. On the other hand, an explanation by borrowing is not completely excluded, because the type of word appearing in the comparisons does not make such a theory improbable, and the shapes of the words lend themselves to it without difficulty.

12. The best attested of the correspondences is *R₁ (Tg. *g*, Jv. *X*, Ml. *r*, NgD *h*, Mlg. *X*). The following are among the best examples:

*baR₁a(h), Tg. *bá:ga*, Jv. *wò-wò*, Ml. *bara*, NgD *bahe*⁷ 'live coal', Mer. *vai-n-áfu* 'ember-of-fire' charcoal.

*baR₁eq, Tg. *bagáq* 'abscess', Jv. *wòh*, Ml. *barah*, NgD *baha*,⁸ Mer. *vái*, *bái* 'sore, abscess'.

*baqeR₁u(h), Tg. *bá:go* 'new', Jv. *wau* 'recently', Ml. *baharu*, *baru* 'new', NgD *báhua* (with metathesis), Mer. *váu* 'new'.

*biR₁aq, Tg. *bigá* 'homalomena', Jv. *wè-wèh-an* name of a water plant, Ml. *birah* 'alocasia', NgD *biha*, Mer. *via*⁹ name of a tuberous plant.

*(dD)aR₁aq, Jv. *rah*, Ml. *darah*, NgD *daha*, Mer. *rá* 'blood'.

*paR₁i(h), Tg. *pá:gi*, Jv. *pe-pe*, Ml. *pari*, NgD *pahi*, Mer. *fái* 'ray (fish)'.

*taR₁uq, Tg. *tá:goq* 'hide', Jv. *tòh* 'wager', Ml. *taroh*¹⁰ 'put, wager', NgD *pa-naho*¹¹ 'offering to the dead', Mer. *táu* 'put down'.

*ZuR₁uq, Tg. *dugóq* 'blood', Jv. *du-doh* 'sap', Ml. *juroh* 'syrup', NgD *joho* 'sap, sauce', Mer. *rú* 'sauce'.

speaking different dialects of Malagasy always succeed in understanding one another, we can conclude that Malagasy is a single language. The term Malagasy can therefore be used in general statements, though in specific citations the name of the dialect will be given (Merina, Sakalava).

⁶ See Dempwolff 2.89, where this reflex is treated as one of the 'unerklärte Ausnahmen'.

⁷ There is no satisfactory explanation of the final *e*.

⁸ We expect a final *e* from **e*.

⁹ Mer. *viha* with the same meaning has an inexplicable *h*.

¹⁰ Colloquial Malay has *tarog* 'put' (with inexplicable *g*) and *taroh* 'wager'.

¹¹ NgD *taroh* 'stake (in betting)' is taken to be a loanword, perhaps from Malay.

13. The single instance of the same correspondence in initial position is the following:

**R₁abut*, Tg. *gá:but* 'tear out', Ml. *rabot* 'tear off', Mer. *dvutra* 'tear out'.¹²

14. The evidence for **R₂* (Tg. *g*, Jv. *X*, Ml. *r*, NgD. *h*, Mlg. *z*) in intervocalic position so far shows it only before **a*:

**baR₂qay*,¹³ Tg. *bagqán*, *bagán* 'molar', Jv. *way* 'jaw', Mer. *vázana* 'molar'.

**beR₂qat*,¹³ Tg. *bigqát*, *bigát*, Jv. *a-bòt*, Ml. *bérat* 'heavy', NgD *behat*, Sakalava¹⁴ *vézatse* 'heaviness'.

**teR₂as*, Tg. *tigás*, Jv. *a-tòs* 'hard', Ml. *těras*, Mer. *téza* 'heartwood'.

**uR₂at*, Tg. *qugát*, Jv. *òt-òt*, Ml. *urat*, NgD *uhat*, Mer. *úzatra* 'sinew, vein'.

15. In initial position, however, the same correspondence is found in one instance before **u*:

**R₂atus*, Tg. *gatos* 'trillion', Jv. *atos*, Ml. *ratos*, Mer. *zátu* 'hundred'.

**R₂umaq*, Jv. *omah*,¹⁵ Ml. *rumah*, NgD *huma* 'house', Mer. *zúma* 'cave', *dzúma-na*¹⁶ 'guardhouse'.

16. Observe that NgD *teras* 'heartwood' and *ratus* 'hundred' (**teR₂as*, **R₂atus*) could be taken together as evidence for a new **R₅* (Tg. *g*, Jv. *X*, Ml. *r*, NgD *r*, Mlg. *z*). If this were done, the evidence of the distinction between **R₁* and **R₂* would be materially reduced, and it would perhaps become possible to find a condition for the appearance of the Mer. *z* from **R₂*. On the other hand this conclusion does not follow immediately, because the remaining instances cannot be subsumed under **R₁* by specifying any condition.

17. The shape of NgD *ratus* 'hundred' offers no difficulty to the theory that it was borrowed from Malay along with many other words, though it could be argued that a word of this meaning was not likely to be borrowed. The shape of NgD *teras* 'heartwood', on the other hand, does offer a difficulty to such a theory. In order to explain this difficulty, as it appears to me, I must first clarify a difference between Dempwolff's approach and mine to the Ngaju-Dayak material.

18. Dempwolff seems to me to have confused the interpretation of the Ngaju-Dayak material by his resort to the theory of an 'older stratum' and a 'younger stratum'. If the features which he assigns to the 'older stratum' be considered marks of inherited words (i.e. as direct reflexes) and the features which he assigns to the 'younger stratum' marks of borrowed words (from Malay and other languages) wherever they contradict those of the first group, the Ngaju-Dayak material can then be treated in the usual way. The resulting condition is that the inherited words (now so considered) generally differ phonetically from a corresponding Malay word, whereas the borrowed words (now so considered)

¹² NgD *má-rabut* 'tear off' is taken to be a loanword, probably from Malay.

¹³ When the Tagalic languages (Tagalog, Bisayan, Bikol) suggest the presence of **q* next to a consonant, I follow the practice of inserting it in the position indicated by Tagalog. No doubts are raised here by Bs., Bk. *báqqay* 'molar'. The comparison of Bs. *búqqat* 'heavy' with **beR₂qat* causes no difficulty, but Bk. *gabát* 'heavy' does—because (besides the metathesis) it has no *q*. It may be that the Bikol word shows the effects of dialect mixture.

¹⁴ See fn. 5.

¹⁵ The explanation of the *o* is not clear.

¹⁶ The appearance of *dz* is the result of analogic change, Dempwolff 2.94.

generally resemble a Malay word with the same meaning. The striking point is that a real difficulty arises in only a small number of instances.¹⁷

19. Now the appearance of *e* in a Ngaju-Dayak word is generally a sign that the word is inherited. The replacement for *ŋ* is generally *NgD a*: *Ml. pēday*, *NgD paday* 'sword'; so also beside *NgD behat* from **beR₂qat* is found *NgD sã-barat* 'as heavy', cf. *Ml. sã-bërat* 'as heavy'. It would then seem likely that *NgD teras* is an inherited word, and if so, its *r* must be considered seriously. However, since *Ml. ẽ* and *NgD e* are probably not far apart phonetically, there is a possibility that the Ngaju-Dayak word is a loanword from Malay, or, if not, a loanword from some other language. These doubts are perhaps sufficient to permit putting aside a single discrepant case.

20. As to the comparisons which seem to lead to **R₂*, the remaining difficulty is the possibility that the difference between **R₁* and **R₂*, namely *Mlg. X* vs. *z*, may be due to a difference in conditions. If so, we cannot point to it now, and hence we must treat **R₂* as a distinctive element; that is, **R₂* reflects either a hitherto unrecognized proto-phoneme or a recognized proto-phoneme plus a hitherto unrecognized second phoneme.

21. The decision between these two possibilities could be perhaps more easily made if the correspondences in the initial position were clear. The difficulties with the reconstruction of **R₂atus* have been discussed above. The reconstruction of **R₂umah* depends on the association of the Merina words; and here a reasonable doubt exists because of their divergent meanings.

22. The preceding discussion can be summed up as follows: (1) **R₁* and **R₂* are distinct correspondences, even though some of the instances here assigned to **R₂* may not belong to it, but (2) it is not clear that the **R₂* correspondence reflects a phoneme that is entirely distinct from **R₁*; it is conceivable that **R₂* is an instance of **R₁* plus an unrecognized second phoneme. The reason for leaving the conclusion in this vague state is the small number of instances.

23. The evidence for **R₃* (*Tg. g*, *Jv. r*, *Ml. r*, *NgD h*, *Mlg. z*) is less simple than that for **R₂*. The instances of the intervocalic correspondence are:

**buR₃ew*,¹⁸ *Tg. bú:gaw* 'put to flight', *Jv. buru* 'pursue', *Ml. buru* 'chase, hunt', *NgD bã-bohaw* 'run away'.

¹⁷ Cf. Dempwolff 2.45 ff., particularly 51 ff. To apply the theory expressed in this paragraph one must divide the *NgD* words cited by Dempwolff into three groups: (1) those that contain a feature of the 'older stratum', (2) those that contain a feature which by Dempwolff's formulas contradicts a feature of the 'older stratum'—hence a feature of the 'younger stratum', and (3) those that contain no distinguishing feature. The last group is assigned by Dempwolff to the 'younger stratum', but its ambiguous position should be recognized. It then turns out that the vast majority of the words in (2) have Malay equivalents which they closely resemble in shape and meaning. There is a residue whose explanation is not obvious; but in the present state of our knowledge of Ngaju-Dayak and of its contacts with other languages, this is not surprising. It follows that Aichele's theory, as quoted by Dempwolff—'dass dort [in Borneo] eine Literatur-Sprache existiert haben muss, ähnlich wie das Kawi auf Java'—is otiose as a way of explaining the 'older stratum'.

¹⁸ If this reconstruction is successful, it is our only example of **-ew*. The reconstruction is an attempt to combine Dempwolff's **buRaw* (based on *TBt. buro* 'pursue', taken with the *Tg.* and *NgD* words) and **buru* (based on *TBt. buru* 'chase', *NgD kã-buro* 'chased away', *Saa huru* 'run', taken with the *Ml.* and *Jv.* words). In reaching my reconstruction I take

**peR₃ges*, Tg. *pigqts*, Jv. *pěřs*, Ml. *pěras*, NgD *pehes* 'squeeze out'.

**qaR₃us*, Tg. *qđ:gos*, Jv. *aros*,¹⁹ Ml. (*h*)*aros* 'current', NgD *asoh*¹⁹ 'flow off'.

24. In none of these three comparisons does a Merina cognate appear; this lack is a basic weakness. The probability that Merina cognates exhibit *z* is built up from a consideration of other correspondences.

25. Another weakness of the theory is that the Javanese words exhibit *r*. In the absence of other evidence we could suppose that the corresponding Malay and Javanese words are connected by borrowing, not by a cognate relation. But it could be posited that a Jv. *ř* in correspondence with Ml. *a* (both regular for **e* before a final consonant) precludes the assumption of borrowing from Malay into Javanese. Jv. *pěřs* then could not be a Malay loanword. There is, however, no reason to exclude the possibility that Javanese borrowed the Malay word before the change of prefinal **e* to Ml. *a*.

26. The following comparison shows Jv. *r* in comparison with Mer. *z*: **iR₃ey*, Jv. *irěy*, Ml. *eray*²⁰ 'black', Mer. *izina* 'dark'. Since the correspondence involves a word from the Javanese 'basic' vocabulary, the importance of the comparison is not inconsiderable. There is little reason to suspect either the Javanese or the Merina word of being borrowed.

27. The conclusion that Jv. *r* in correspondence with Malagasy *z* is an instance of **R₃* depends essentially on the following two comparisons involving respectively the morpheme-final and the morpheme-initial reflexes:

**aluR₃*, Tg. *qđ:log* 'ford', Jv. *alor* 'morass', Ml. *alor* 'channel', NgD *l-aloh-an* 'navigable water', Mer. *dlu* 'pour', *alúz-ina* 'what is to be watered'.

**R₃akit*, Jv. *raket* 'pair, couple', Ml. *raket* 'laid parallel, raft', NgD *hakit*, Mer. *zđhitra* 'raft'.

28. The words here taken to be cognate appear to belong together. The meanings do not suggest that the Javanese word is borrowed from Malay. In each case the Merina word has *z*. These sets, together with the set under **iR₃ey* above, can be satisfied only by the reconstruction of **R₃*. A counter-argument could

TBt. *buru* as probably a loanword and NgD *-buro* as a Malay loanword (leaving the final *o* unexplained) which has acquired a NgD prefix. The reconstruction **-ew* is not interdependent with that of **R₃*; the latter depends on the assumption that Jv. *buru* is inherited and not a loan from Malay. In either case, however, we can reach the reconstruction of **-ew*. The reflexes assigned to **-ew* (Tg. *-aw*, TBt. *-o* (?), Jv. *-u*, Ml. *-u*, NgD *-aw*) are parallel to those assigned to **-ey* (Tg. *-ay*, TBt. *-e*, Jv. *i*, Ml. *i*, NgD *-ey*, Mer. *-i*); e.g. **qatey*, Tg. *qatáy*, TBt. *ate-ate*, Jv. *ati*, Ml. (*h*)*ati* 'liver', NgD *atey* 'feelings', Mer. *ati* 'liver'. Now Saa has *-e* from **-ey*, as in *s-ae* 'liver, feelings', and we should perhaps expect Saa *-o* from **-ew* rather than the *-u* of *huru*; but the discrepancy is explained as assimilation. The interesting part of this Saa word is its *r*; for Saa *l* reflects **R₁* (**paR₁i*, Saa *heli* 'ray') and **R₂* (**uR₂at*, Saa *ule-ule* 'tendon, vein'). If it is true that Dempwolff's two reconstructions above can be replaced by one, and if the association of Saa *huru* is correct, we have further evidence of a distinction other than **R₁* and **R₂*.

¹⁹ This is a combination of Dempwolff's **haRus* and **qaRus*; the initial reflexes are dealt with in my forthcoming monograph on the PMP laryngeals. Jv. *đs-đs* 'hiss', cited by Dempwolff under his **haRus*, is discarded because of its deviant meaning. NgD *hđrus-an* (as well as its by-form with metathesis, *rahus-an*) 'river-bed' is taken to be or contain a loanword, perhaps from Malay.

²⁰ There is no satisfactory explanation of Ml. *e* from **i*.

assume that the Javanese words in all the instances assigned to * R_3 are ancient loanwords (ancient so as to allow for semantic change) from a language like Malay in which * R_2 (or, where the Malagasy cognate is lacking, possibly also * R_1) became r . Even the appearance of the Mer. z in * $aluR_3$ is suspect, since it occurs in a position exposed to the workings of analogic change.²¹

29. The evidence for * R_4 (Tg. g , Jv., Ml., NgD, Mlg. r) is simple:

* $baR_4u(h)$, Tg. *bali-bago*, Jv. *waru*, Ml. *baru*, NgD *baro*, Mer. *váru*, *báru* 'hibiscus'.

* $habaR_4at$, Tg. *habá:gat* 'west or south wind', Jv. *barat* 'west monsoon', Ml. *barat* 'west', NgD *barat* 'west, west wind, storm (from any direction)', Mer. *aváratra* 'north'.

* uR_4ita , Tg. *p-ugi:taq*, Jv. *k-ëritò*, *g-ëritò*, Ml. *g-urita* 'polyp', Mer. *h-uríta* 'squid'.

30. No instances of * R_4 in morpheme-initial position are citable. Several possible instances occur in morpheme-final position, but here evidence is lacking from Malagasy, which lost its final consonants. A likely instance is the following:

* $timuR_4$, Tg. *tí:mog* 'south', Jv., Ml., NgD *timor* 'east', Mer. *a-tímu* 'south'.

31. In respect to * R_4 the evidence depends not alone on Javanese, but also on the agreement of Ngaju-Dayak and Merina in showing r . The only possible escape from the conclusion reached here is to say that the Merina (as well as the Javanese and Ngaju-Dayak) words are borrowings from a language which had r regularly in these words.

32. This is in effect the explanation given by Dahl.²² He finds that Malagasy is most closely related with Maanyan (southeastern Borneo), and sets the probable date of the departure of the Malagasy circa 400 A.D. He then assumes that these words were borrowed by the parent language of these two languages from a language in which * R became r .

33. The three comparisons of the intervocalic correspondence undoubtedly reflect a PMP word. With * $baR_4u(h)$ we can compare Fi. *vau*, To. Fu. Sm. *fau* 'hibiscus'; with * $habaR_4at$, Fi. *z-avā* 'storm', To. *afaa* 'storm', *afats-ia* 'stormed over', Fu. *afa-qa* 'storm', Sm. *afaa* 'storm', *afat-ia* 'stormed over'; with * uR_4ita , Fi. *k-uita* 'polyp'. A theory that borrowing occurred in all cases could be viewed as an avoidance of the reconstruction of a new phoneme. Though it can be granted that the hypothesis of borrowing introduces no complications under Dahl's theory, there appears to be at least an equal (if not better) chance that we are dealing here with inherited words.

34. The Maanyan cognates found by Dahl of words here referred to * R_4 are Mny. *barat* 'storm, west' and *timur* 'east'. These words are thus the sole Maanyan support for his theory of borrowing, but they equally well support a theory of a separate * R_4 .

35. The Maanyan reflex of both * R_1 and * R_2 is X (no instances of a word presumably containing R_3 being available): * $baqeR_1u(h)$, Mny. *wao* 'new'; *(dD) aR_1aq , Mny. *i-ra* 'blood'; * beR_2qat , Mny. *weat* 'heavy'; * teR_2as , Mny.

²¹ Cf. Mer. *tavéz-ina* 'to be made savory' beside *tavi* 'fat, savory' (from *(tT) $abe(q0)$, Tg. *tabáq*, TBt. *tabo* 'fat, savory'), in which the z must be analogical in origin.

²² *Malgache et maanyan* 59.

teah 'hard'; **uR₂at*, Mny. *uwat* 'vein'. If there was a separate Maanyan-Malagasy unity, it evidently had **R₁* and **R₂* as distinct phonemes.

36. Another point of interest is the Trukese evidence, which points to a cleavage between **R₁*-**R₄* on one side and **R₂*-**R₃* on the other. Tr. *jafar* 'shoulder' compares well with Tg. *qabá:ga*, NgD *baha* 'shoulder', and the NgD *h* will satisfy all but **R₄*. On the other hand Tr. *k-yys* 'squid' compares well with the words grouped under *-*uR₄ita*, and the second part of *sini-fē* 'bark-hibiscus = hibiscus' compares well with the words grouped under **baR₄u(h)*. We could conclude that **R₄* was lost in Trukese, whereas at least one of the others became Tr. *r*. Which one it is that appears in the word for 'shoulder' can be determined by citing Tr. *mine-fē* 'new' (cf. *mine-nnöm* 'old'), which compares well with **baqeR₁u(h)*; and, as a contrast with *jafar*, Tr. *cca* 'blood', which compares well with *(*dD*)*aR₁aq*. It then follows that neither **R₁* nor **R₄* occurs in the word for 'shoulder'. We can perhaps go a step further, and associate Tr. *amürenipwin*²³ 'dew-night = dew' with Tg. *hamóg*, Jv. *l-amor* 'dew'; this could be an instance of **R₃*. Similarly we can cite Tr. *jegē-rēw* 'one thousand' as matching **R₂atus*. We appear to reach the hypothesis that both **R₂* and **R₃* became Tr. *r*, whereas *R₁* and *R₄* disappeared.²⁴

37. I cannot help feeling that the evidence presented here is not entirely solid. The arrangement of the discrepant correspondences suggests that hitherto unrecognized distinctions may be involved. Much work of the same kind remains to be done with regard to other discrepancies in Dempwolff's work. The conclusion is inevitable that Dempwolff's procedure was insufficient because it omitted the third step: it failed to re-evaluate the results reached by the deductive second step. The procedure of excluding the evidence of a language from consideration must be equally applied; and this is achieved by neglecting reflexes under a theory of analogic change and borrowing.

38. When, as here, there is a considerable doubt as to the explanation of a given correspondence, we may resort to 'problematic' reconstructions, leaving a final solution to the future. Now that the like correspondences (deviant in terms of Dempwolff's **R*) have been placed together, it will be much easier to evaluate new evidence. In any case, by following this procedure we can maintain a principle of 'total accountability' with respect to recurrent correspondences, without being forced prematurely into the reconstruction of new proto-phonemes.

²³ Cited directly from S. Elbert, *Trukese-English and English-Trukese dictionary* (1947). Otherwise Trukese words are cited in the same way as in Dyen, *On the history of the Trukese vowels*, *Lg.* 25.420-36 (1949).

²⁴ If Saa *r* reflects **R₃* (see fn. 18 end), then Saa *ahala*- 'shoulder' indicates **R₂* in the PMP reconstruction.

TWO TONE PATTERNS IN TANGSIC

GEORGE A. KENNEDY

Yale University

Possibly the most widely disseminated fact about the Chinese language is that each syllable has a definite musical accent or tone forming an inalienable part. This fact emerges from even the most casual inspection of a Chinese dictionary, in which the syllabic symbols are each labelled with one or more tones, and where it is obvious that the intelligibility of a particular syllable depends on its having the correct tone. Persons who try to talk standard Chinese have to be told that a large number of these tones disappear in connected speech, which only goes to show the universal sloppiness of people. A pedantic speaker, who is at the same time literate, can restore the tones by visualizing the graphs with which the speech would be written down. But it must be concluded that for illiterates the intelligibility of their communications cannot depend on a theoretical knowledge of the tones for syllables that have no tones. An examination of Tangsic brings to light in that Wu dialect a situation rather different from what is ordinarily imagined. (For a definition and general description of the Tangsic dialect see *Lg.* 28.457-8.)

This paper is not concerned with a phonetic description of the dialect except as it relates specifically to the topic. These introductory remarks give only what is necessary for reading the spelled words with some approximation to the originals. Syllables may begin with the following sounds:

			1	2	3	4	5	6	7
a	VOICELESS	unaspirated	<i>p</i>	<i>t</i>	<i>tz</i>	<i>k</i>	<i>ʔ</i>		
b		aspirated	<i>ph</i>	<i>th</i>	<i>ts</i>	<i>kh</i>	<i>h</i>	<i>f</i>	<i>s</i>
<hr/>									
c	VOICED	non-nasal	<i>bh</i>	<i>dh</i>	<i>(dz)</i>	<i>gh</i>	<i>hh</i>		
d		nasal	<i>m</i>	<i>n</i>	<i>l</i>	<i>ng</i>	<i>(nh)</i>	<i>v</i>	<i>z</i>

Syllables may be closed by *-n* or *-ng*; these are in complementary distribution and only weakly differentiated, but for comparative reasons are best written as two different elements. Syllables may also have a glottal closure, symbolized by *-q*. In closed syllables only five vowel sounds appear, the last being extremely rare: *-a-*, *-e-* [ə], *-i-*, *-o-*, *-u-* [ʊ]. The vowel sounds in open syllables are more varied, and may be written in part with digraphs: *-i-*, *-ea* [ɛ], *-ae* [æ], *-eo* or *-oe* [ō], *-a-*, *-u-*, *-ao* [o/ə], *-oa* [ɔ]. In addition there are the pseudo-diphthongs *-ie* [i·ɪ/i·ə] and *-ue* [u·ʊ/u·ə]. Of the initial sounds, columns 3 and 7 and all of row d may be fully syllabic without any vowel. Between the initial and the head vowel may appear a medial *y* or *w*. These have no effect on the surroundings, except that *ky-*, *khy-*, *ghy-*, *hy-* are [tʃ, tʃʰ, dʒ, ʃ].

A measured analysis of individual tones is similarly outside the scope of this paper, since we are concerned with things that occur in synthesis. Absolute pitch has no bearing on the subject, but a tone will appear in relatively different keys,

or registers, according as it accompanies a syllable with a voiceless or with a voiced initial. The interval between these keys is usually five to seven semitones on the chromatic scale. If (to make an analogy) a melody with voiceless initial is played in the key of F, then the same melody with voiced initial is likely to be played in the lower key of C. In order to describe the tones, which are analogous to the melodies, without a complicated apparatus of diagrams, we may adopt a simplification of the picture in terms of numbers. Let the numbers 1 to 5 represent five points in rising order of pitch. The precise intervals between these points are immaterial, so long as 2 is somewhat lower than 3, 5 is much higher than 1, and so on. Then the notation 24 stands for a continuous rise in pitch between the points indicated, 51 represents a sharp fall, while 33 marks a level tone with unchanging pitch. These tones may be played in any key in accordance with the quality of the speaker's voice or his emotion of the moment. Two contrasting keys are used by each speaker, paralleling the difference in type of initial for particular syllables, and this may be shown by using boldface numerals for the lower key. With this rough symbolism established, we can state the general correspondence of Tangsicc tones on monosyllables to the first three historical categories:

	UPPER KEY	LOWER KEY
I	33 'level'	24 'low rising'
II	51 'high falling'	51 'low falling'
III	24 'high rising'	24 'low rising'

The fourth historical category is represented in Tangsicc by closed syllables ending in *-q* (glottal stop). The duration of these syllables is so short that neither continuity nor change in pitch is apparent, and it has often been questioned whether there is any justification for regarding IV as a tonal category. For the time being we regard them simply as 'syllables in *-q*'. In addition to the symbolism by numbers we may use diacritics when convenient according to the scheme *lèvèl, rising, fàlling*.

Textbooks on standard Chinese pay attention to two factors that are viewed as producing changes in the 'inherent' tones of syllables. Sandhi is a mechanical change arising from the juxtaposition of certain tones; the neutral tone is a disappearance of characteristic tone resulting from complete lack of stress. Following this line of interpretation we must begin for Tangsicc with monosyllabic words, as the only syllables for which the inherent tone can be directly known.

PATTERN A, 'HIGH MOUNTAIN'

kāo 33 'high', *sāe* 33 'mountain'; *kāo-sāe* 33-33 'high mountain'. The only change to be noted is a closer juncture between the syllables, which, so far as the tones are concerned, makes them flow in a continuous line without hiatus. Musically speaking the effect of the juncture is precisely like that of the tie between two quarter-notes on the same line of the staff, calling for the production of one half-note.

kà 51 'artificial', *sāe* 33 'mountain'; *kà-sāe* 53-31 'rockery'. The 51 of the first syllable is cut off, while the 33 changes to 31. This phenomenon might be de-

scribed as the second syllable taking the tone of the first, or as the falling tone of *kà* inducing a falling tone on *sāe*. Persons familiar with Pekingese tonal behavior may compare this with the neutral tone after a falling tone, which in effect is a tag end of it. The tones of the two Tangsic syllables again fall in a continuous descent without hiatus.

dhú 24 'large', *sāe* 33 'mountain'; *dhū-sāe* 22-44 or *dhū-sāe* 22-24 'large mountain'. The second variety is associated in the speaker's mind with slow or effeminate speech. In either case 24 has become 22, a low tone without rise which is not one of the original tones. This change is comparable to the change of the 3rd Pekingese tone before another syllable with different tone. In the effeminate variety the second syllable could again be described as taking the tone of the first; the two tones again form a continuous curve without hiatus. In the 'masculine' variety (which is not confined to the speech of males) there is a break, and *sāe* is level, though on a lower pitch than when in isolation. Note that 44 of the lower key is lower than 33 of the higher key.

The changes in tone have already begun to appear complicated, but one simple statement will cover all four cases shown: THE ORIGINAL TONE OF THE FIRST SYLLABLE IS SPREAD OVER THE TWO COMBINED SYLLABLES. This is the same as saying that two syllables are embraced under one tone. An extension of the number symbolism is then suggested, as follows: 33 + 33 > 3003. The notation 3003 denotes an expanded 33, with zeros to fill out the number of syllables involved. In the same way 51 + 33 > 5001, and 24 + 33 > 2004; the latter defines either 22-44 or 22-24. This interpretation of the tonal change may now be tested by other examples:

tzōng 33 'middle', *nytn* 24 'man'; *tzōng-nytn* 3003 'middleman'
syào 51 'small', *nytn* 24 'man'; *syào-nytn* 5001 'children'
dhú 24 'large', *nytn* 24 'man'; *dhū-nytn/nytn* 2004 'adults'
pēn 33 'ice', *sēa* 51 'water'; *pēn-sēa* 3003 'ice-water'
kwèn 51 'boiling', *sēa* 'water'; *kwèn-sēa* 5001 'boiling water'
tséo 24 'smelly', *sēa* 'water'; *tsēo-sēa/sēa* 2004 'smelly water (disinfectant)'

The additional examples all conform to the statement.

There are two observations to be made on the behavior of tone as so far exhibited in Tangsic that must be shocking to the comfortable conventional view. Given only the expression *syào-nytn* 5001 'children', it is impossible to deduce the 'inherent' tone of *nytn*, nor is the situation helped by bringing in *tzōng-nytn* 3003 and *dhū-nytn* 2004. The syllable *nytn* with its inherent tone, as required for a dictionary, can be found only if *nytn* occurs as an independent word, or in another pattern to which this kind of tonal behavior does not apply. For the illustrations used, as already mentioned, I deliberately chose syllables with independent meaning. In the second place, if a phrase of this pattern commences, for example, with *kwèn* 51 'boiling', the inherent tone of the word for what it is that boils is quite irrelevant. Anything at all described as 'boiling' will, in that condition, have a falling tone, whatever it may have otherwise. And if one can divorce oneself completely from the dictionary, he will have to conclude that the tone on X at any particular moment is wholly determined by the tone of

the attribute. Then if two words like *sāe* 'mountain' and *sāe* 'umbrella', distinguished by tone, are both described by the attribute 'large', their separate identity disappears, and both expressions become *dhū sāe* or *dhū sāe* 2004. It is, as a matter of fact, impossible for the Tangsite to know without additional context whether this means 'a large mountain' or 'a large umbrella'. The effect, then, of spreading the tone of *dhū* over a following syllable *X* is to announce that the *X* is described by *dhū*, but at the same time to confuse the identity of *X*. The intelligibility of *dhū-X* is now dependent not on the recognition of *X*, but on the recognition of *dhū-X* as a whole.

Before involving ourselves further in such speculative thought, we should take note of the behavior of syllables in *-q*.

bhaq 'white', *sāe* 'mountain'; *bhas-sāe* 'white mountain'; similarly *bha-nyin* 'white man', *bhas-sāe* 'white water'. First to be noted is the disappearance of the *-q*, or its assimilation to the initial of the following syllable; and second, that the falling tone on the second syllable begins somewhat higher and falls much more abruptly than the tone that has been symbolized as 51. We may denote it as 60. To relate this tonal change to the general principle that has been proposed may seem fanciful, but the description of Tangsic will be simpler if a connection can be made. The total tone of 'white mountain' is discontinuous, and may properly be symbolized as 3600. If this represents the spread of the tone of the first syllable, the tone on *bhaq* must be something of the same nature. But, as I have noted, syllables in *-q* are so short that no tonal pattern is distinguishable to the ear. The distinctive feature of the first syllable that can be assumed to have been lost or passed on is the glottal closure, and this appears to be reflected in a particularly sharp falling tone. It is not difficult, if one wishes, to relate these two things phonetically as being both a cutting off of sound. For a single syllable the cutting off is achieved by an abrupt closure of the glottis, while for a dissyllable it is achieved by an abrupt fall of tone to zero on the second syllable.

In the combination *bha-nyin* 'white man', despite the loss of *-q*, the first syllable remains distinct from an original *bha* through having a shorter vowel. But the more important distinction arises from the fact that no *bhá* or *bhà* could produce the total pattern 3600: the intelligibility of the expression depends, as in previous cases, on that total pattern.

There remains the case of a syllable in *-q* as the second element: *kwoq* 'country', *yīn-kwō(q)* 33-3 'England', *ngū-kwoq* 22-4 or *ngū-kwō* 22-24 'Russia', *mèa-kwoq* 55-1 or *mèa-kwō* 53-31 'America', *fak-kwoq* 2-5 or *fak-kwō* 3-60 'France'. The single numeral is intended to symbolize a pitch without appreciable duration, associated with glottal closure. This closure may be absent in the second syllable, in which case the shape of the combination follows the patterns already described. The alternative forms here show more clearly than ever that the essential principle in combination is to extend the initial tone.

It may be concluded that all of the combinations so far examined represent the extensions of four tonal shapes: level, rising, falling, and abrupt; for the last one we add the circumflex to the diacritics already chosen. Then the four distinctive tones on monosyllables are shown by *kāo* 'high', *dhū* 'large', *syào* 'small', *bhāq* 'white'. If it were typographically feasible, an elongated writing of these

diacritics over two syllables together would quite accurately represent the tonal shape of dissyllables in pattern A. Our present concern, however, is not practical orthography but rather the bearing of this fact on the units of Tangsic speech. Even the most conservative among Chinese scholars might admit that some of the listed combinations deserve to be called 'words'. But if the term word is applied to any of them, it must be applied to all, since there is no formal difference between 'high mountain' and 'England'.

The alternative is to treat syllables in the conventional way as individualized units which, to be utterable, must have inherent tones, and then to state the rules of tonal sandhi in composition. When one proceeds from a text written, or at least writable, in Chinese graphs, the supposedly inherent tones will be found attached to the graphs in a dictionary. But Tangsic, happily or unhappily, has never been written with Chinese graphs, and to a considerable extent cannot be written with the available stock. The lack of this crutch often makes the individualizing of syllables difficult or even impossible. To the first syllable in pattern A the linguist, by deduction from the total shape, can assign a tone, though not necessarily a meaning. Thus from 'England' he can abstract the syllable *yīn*, which to the Tangsite is meaningless except as 'the first syllable of England'. From 'Russia' can be taken *ngú*, which means 'goose' or 'to be hungry'. The *mèa* of 'America' is again meaningless, while *fāq* can be understood as 'the rising of dough'. In a second syllable the form can be identified if it ends in *-q*; otherwise, as already noted, its tone cannot be determined except by other means.

A common noun suffix in Tangsic is shown by the following names of fruits: *āng-tz* 'cherry', *dháo-tz* 'peach', *tzào-tz* 'date', *kyít-tz* 'orange'. The tones on *tz* are successively 33, 44/24, 31, and 60; and since it is never anything but a suffix, it is quite impossible to determine what inherent tone it might theoretically have. This problem is easily solved, by Chinese and western scholars alike, by pointing to a graph—a natural action, since the problem is discussed only by literate people. Even Tangsites sometimes become literate by learning to read Mandarin, which though a different dialect contains such words as *jjutz* 'orange'. Recognizing this to refer to the same thing as Tangsic *kyít-tz*, the Tangsite sees that the graphs used to write *jjutz* in Mandarin can quite comfortably be borrowed and read *kyít-tz*; and as soon as a graphic representation is provided for the Tangsic suffix, its inherent tone is settled. This is without doubt all very proper in a historical way, since the two words must have the same parentage. But the point here is that the conclusion for Tangsic could be reached only in a devious and sophisticated manner, and that within the limits of the dialect spoken by the illiterate Tangsite the notion of an inherent tone on *tz* is meaningless.

PATTERN B, 'COOK SOUP'

The word for 'cook' or 'burn' is *sāo*, and the word for 'soup' is *thāng*. In the expression 'to cook soup' both tones remain level, but instead of forming a continuous note as in *kāo-sāe* 'high mountain' they show a break. The pitch of *thāng* is 33, as in isolation; but the pitch of *sāo* is dropped a half or a whole note.

Three other words are *má* 'sell', *mà* 'buy', and *khyúq* 'eat' or 'drink'. The expressions 'sell soup', 'buy soup', and 'eat soup' all have precisely the same tonal pattern as 'cook soup'. *Thāng* 'soup' remains on a level 33, while the preceding words all become level on a slightly lower pitch. We shall denote the latter tone by S. The -q may be retained in careful speech, but commonly disappears. There is never any assimilation to the following consonant.

A complete chart of this pattern may be given in tabular form:

	<i>thāng</i> 'soup'	<i>mīe</i> 'noodles'	<i>tzəo</i> 'wine'	<i>tzôq</i> 'porridge'
<i>sāo</i> 'cook'	S-33	S-24	S-51	S-3
<i>má</i> 'sell'	S-33	S-24	S-51	S-3
<i>mà</i> 'buy'	S-33	S-24	S-51	S-3
<i>khyúq</i> 'eat/drink'	S-33	S-24	S-51	S-3

In all 16 combinations the second syllable retains its original tone, while the first syllable becomes S—that is, a level tone lower than the usual 33.

To those accustomed to the concept of inherent tone, this pattern may seem even more extraordinary than pattern A: except for syllables in -q, there is no way to determine from the dissyllabic expression what tone belongs to the first element. This results, as the examples show, in the curious situation that whereas 'buy' and 'sell' are different words in isolation, 'to buy anything' becomes indistinguishable from 'to sell anything'. The Tangsite cannot, in fact, tell without further context whether *ma-tzəo* means 'buy wine' or 'sell wine'. In the second place, the tone on the first syllable in pattern B is not a sandhi change, in the sense of being dependent on the nature of the preceding or following tone. Its character is S, regardless of what follows. Third, if this character S is viewed as a change from an original inherent tone, the change is obviously anticipatory; but then the question becomes, 'anticipatory to what?'

Comparison of patterns A and B shows that the tonal behavior cannot be explained in any purely mechanical way. Within each pattern, it is true, the mechanical condition of stress can be accepted as determinative. It is phonetically demonstrable that in A the louder stress is on the first syllable, in B on the second. It is sufficient to say, then, that a stressed syllable extends its tone over all unstressed syllables that follow in close juncture; and that an unstressed syllable in close juncture before the stressed syllable takes a special tone, here symbolized by S, which may be described as a suspense tone. Adequate though this statement is to describe the phenomena noted, it does not, of course, define the conditions that determine which syllable is to be stressed. Over and above the inherent tones and the mechanical adjustments accompanying lack of stress there is some other factor; and this factor cannot be anything less than morphological. It will be clear from the following examples that combinations of the same semantic roots express a syntactic difference by different tonal patterns which, in the last analysis, can be described as different stress patterns.

ʔéa 'love', *nyŋn* 'man': *ʔéa-nyŋn* 2004 'sweetheart (loved person)' (A),

ʔea-nyŋn S-24 'to love people' (B)

- sāo* 'cook', *tzəo* 'wine': *sāo-tzəo* 3003 'distilled spirits, Branntwein' (A),
sāo-tzəo S-51 'to cook wine (improbably)' (B)
yáo 'row', *záo* 'boat': *yáo-záo* 2004 'a rowboat' (A), *yáo-záo* S-24 'to row a
 boat' (B)
tsáo 'fry', *váo* 'rice': *tsáo-váo* 5001 'fried rice' (A), *tsáo-váo* S-24 'to fry
 rice' (B)

The same behavior is shown somewhat more complexly in the following: *khyw-váo-səo* 36001 'the right hand (eating hand)' (A), *khyú(q)-váo* S-24 'eat rice (or food in general)' (B)

When the verb-object expression (B) itself becomes a modifier of something else, the whole pattern is reshuffled and takes off from the first syllable to form pattern A:

- dhó(q)-s* S-33 'to read books, to study' (B), *dho-s-nyn* 36000 'a scholar' (A)
pháo-ghyéu S-24 'to bat balls, to play tennis' (B), *pháo-ghyéu-záng* 30003 'a tennis court' (A)
bháo-móa S-51 'to race horses' (B), *bháo-móa-thén* 20004 'a racecourse' (A)

Extensions over four syllables are illustrated by the following:

- kyi-máo-táo-tz* 300003 'feather-duster'
yáng-thi-kwúe-dhéu 200004 'tin can'
páo-hyé-kóng-s 500001 'insurance company'
káp-pi-nyn-ká 360000 'next-door neighbors'

The scholar in Chinese will recognize that the speculations to which this brief note on Tangsic might lead are much too fluid to be contained within a short article. It is beyond question that tone is word-distinctive on monosyllables in Tangsic. But these monosyllables form a very small percentage of the vocabulary. Most of the time the tonal pattern is functioning morphologically. Given only two occurrences of the syllable *tsáo* as illustrated in the last paragraph, one would reach the conclusion that there was in Tangsic an adjective *tsáo* 'fried' and a verb *tsáo* 'to fry'. Then the tone on monosyllables is in such a case also morphological. And one might be led to the broad guess that the function of tone in Tangsic is essentially not to distinguish otherwise homophonous syllables, but to express syntactic relationships. From spoken material alone and without reference to historical dictionaries it is extraordinarily difficult, as I have often discovered, to make up a word-list for Tangsic in terms of monosyllables with fixed tones. There is a limited number of nouns and verbs that can be given with tones in isolated form; but for a large part of the vocabulary only an appeal to a hypothetical system of writing can settle the question of tone. It has always been hard for sinologists not to feel that the written graphs are in fact the final authority. In a realistic treatment of Tangsic they cannot be.

ON THE TONE SYSTEM OF THE MIAO-YAO LANGUAGES

KUN CHANG

University of Washington

The Miao-Yao languages are spoken by tribes scattered among the mountainous regions of southwestern China (in the provinces of Hunan, Kueichow, Kuangtung, Kuanghsi, Yünnan, Szuchuan), in Viet Nam (in the area of Tonkin), and in Thailand.¹ These tribes bear various Chinese names, with either *Miao* or *Yao* as their last element. On the basis of the information that I now have, I have tentatively classified this group of dialects into two main divisions: the first preserves the syllable-finals *-p*, *-t*, *(-k)*, *-m*, the second has lost them. The dialects of the second division in turn form two subdivisions: one displays a highly complicated system of initials, the other a comparatively simple one. In this paper ten dialects are discussed:² MKC (Hungchan Miao in Kaop'o, Kueichu, Kueichow), MKS (Yi Miao in Kechengchai, Kuangshun, Kueichow), YLP (Hei Yao in Yaolu, Lipo, Kueichow), YYT (Tahua Yao in Hsishankai, Yungts'ung, Kueichow), MTK (Hei Miao in Shihtungk'ou, Taikung, Kueichow), MCF (Hei Miao in Chengfeng, Kueichow), MJC (Hei Miao in Kaot'ungchai, Jungchiang, Kueichow), YYL (Yao in Yaoling, in the districts of Lienchow, Lienshan, and Yangshan in Kuangtung), YHN (Ch'ing-yi Yao in Haining, Tonkin, Viet Nam), YTP (Tapan Yao in Tonkin, Viet Nam). Most of these dialects belong to the second main division; MKC, MKS, and YLP differ from MTK, MCF, MJC, and YYT by having nasal-prefixed initials (stops and affricates) and initial consonant clusters with *-l-* or *-r-* as the second element. The dialects of the first main division are YYL, YHN, and YTP; they are cited only to show the distinctive finals *-p*, *-t*, *(-k)*, *-m*. The many seemingly irreconcilable irregularities in the phonology of the three Yao dialects YYL, YHN, and YTP may be due to the inaccuracy of the records made by Wong and Savina.

An extensive study of cognates among the seven best-known dialects shows

¹ Bernhard Karlgren, *Études sur la phonologie chinoise*, *Archives d'études orientales*, Vol. 15 (1915-26); Henri Maspero, *Contribution à l'étude du système phonétique des langues Thai*, *BEFEO* 11.153-69 (1911); Fangkuei Li, *The distribution of initials and tones in the Sui language*, *Lg.* 24.160-7, (1948); *Tones in the riming system of the Sui language*, *Word* 5.262-7 (1949); *The Mak language*, *Bulletin of the Institute of History and Philology, Academia Sinica* 19.1-80 (1949).

² The material on dialects MKC, MKS, YLP, YYT, and MJC was gathered during my field trip in 1941 and 1942 under the auspices of Academia Sinica. The data on MTK were kindly furnished by Fangkuei Li. The examples of MCF, YYL, YHN, and YTP are taken from the following publications: Joseph Esquirol, *Dictionnaire kanao-français et français-kanao* (Hong Kong, 1931); Sikling Wong, *Phonetics and phonology of the Yao language, description of the Yaoling dialect*, *Lingnan science journal* 18.425-55 (1939); F. M. Savina, *Dictionnaire français-man* (Hanoi, 1926). The publications on the Miao dialects in Tonkin by Savina, on the Ch'ing Miao dialect in Eshan (Yünnan) by Huanien Kao, and on other Miao dialects by Yu Wen and others are a bit sketchy for comparative study. This paper has been read and criticized by Bernard Bloch, Y. R. Chao, Loren W. Fessler, and Fangkuei Li. A similar article by me, in Chinese, was published in *Bulletin of the Institute of History and Philology, Academia Sinica* 16.93-110 (1947).

eight correspondences of tones. The number of tones in a single dialect is six (MKC), seven (MKS, YLP), or eight (YYT, MTK, MCF, MJC). Some dialects (YYT, MTK) have five level tones of different heights besides rising tones and falling tones.³

	I	II	III	IV	V	VI	VII	VIII
MKC	¹ high rising	² high level	³ low rising	⁴ low falling	⁵ high falling	⁶ mid level	⁷ high falling	⁸ high level
MKS	¹ mid high level	² high level	³ high rising	⁴ mid falling	⁵ mid level	⁶ mid low level	⁷ mid level	⁸ low falling
YLP	¹ mid low level	² high level	³ rising	⁴ high falling	⁵ mid level	⁶ low level	⁷ high level	⁸ low falling
YYT	¹ rising	² mid level	³ mid low level	⁴ low level	⁵ high level	⁶ mid high level	⁷ high falling	⁸ low falling
MTK	¹ mid high level	² high falling	³ high level	⁴ mid low level	⁵ high rising	⁶ mid level	⁷ low rising	⁸ low level
MCF ⁴	¹ mid one	² high one	³ high two	⁴ mid two	⁵ mid low	⁶ low	⁷ high short	⁸ low short
MJC	¹ mid high level	² high falling	³ high level	⁴ low falling	⁵ high rising	⁶ mid level	⁷ low rising	⁸ low level

The last two categories of tones are supposedly those of words originally ending with *-p*, *-t*, (*-k*), which are still preserved in YYL, YHN, and YTP.

	'duck' ⁵	'drink'	'bear (animal)'	'see'	'tongue'
MKC	'ə ⁵	he ⁵	tle ⁵	pe ²	mple ²
MKS	'ău ⁵	hu ⁵	qlai ⁵	pău ⁷	mplai ⁷
YLP	'au ²	hi ²	tljăi ²	pău ⁷	ntljăi ⁷

³ The terms 'high', 'mid', 'low' used in the following table refer to relative height within each dialect. The high rising tone of MKC does not necessarily have the same contour as the high rising tone of MKS. Understanding now that the pitch of the tones is taken in a doubly or triply relative sense, it might be more precise and clear, as suggested by Y. R. Chao, to give a graphic representation. But owing to typographical difficulties, I regretfully sacrifice the beautiful chart made for this paper by Chao.

⁴ The values of the tones in MCF are not clearly described by Esquirol.

⁵ Cf. Ancient Chinese *ăp 'duck'. This paper does not treat borrowings from Chinese, which have presumably taken place at various periods. Since the history of the cultural contact between the two peoples has not been well studied, it is difficult to say anything definite about the Chinese loanwords in the Miao-Yao languages.

Notes on the symbols used in this paper: ə neutral vowel; ı unrounded u; ă short back a; ɔ̌, ɛ̌ domals; ť, ď, ɲ̌ palatals; ɲ̌ back nasal; G, Ň velars; ɲ̌, ɲ̌, ɲ̌ voiceless nasals; ľ voiceless lateral; ʔ glottal stop.

⁶ I can offer no explanation of the initial *ph*- in the examples taken from Savina's *Dictionnaire*.

⁷ YYT has nasal initials as counterparts to the nasal-prefixed initials in MKC, MKS, YLP, in words of both tone types.

⁸ I admit that there is a choice between regarding tone or voicing as phonemic.

YYT	'ae ⁷	ho ⁷		pa ⁸	mai ⁸
MTK		hau ⁷		pa ⁸	ñe ⁸
MCF		hau ⁷	hi ⁷	pañ ⁸	ñi ⁸
MJC	'au ⁷	hu ⁷	tje ⁷	pu ⁸	ñe ⁸
YYL	'ap	hup		bat	bet
YHN ⁸	'ap	hop	kiop	phôt	biet
YTP ⁸	'ap	hop	kiôp	phôt	biet

A scrutiny of the distribution of tones and initials in cognate words shows that the odd-numbered tones and the even-numbered ones constitute two different types, in that certain initials occur only with tones of the first type. These initials are: the glottal stop; the voiceless (or aspirated) nasals *mh-*, *nh-*, *ñh-*; voiceless *lh-* in MKC, MKS, YLP, and YYT, or aspirated voiceless *lh-* in MTK and MCF; aspirated voiceless stops and affricates; the nasal-prefixed aspirated voiceless stops and affricates of MKC, MKS, and YLP, which correspond to plain voiceless stops, affricates, and spirants in MTK, MKS, and MJC. All these initials can be easily reconstructed for Pre-Miao-Yao as glottal stop or aspirated voiceless consonants. The other initials, appearing in words of both tone types, probably have two different origins: voiceless in words with the first tone type, voiced in words with the second tone type.

The correspondences between the nasal-prefixed initials of MKC, MKS, and YLP and their parallels in MTK and MJC⁷ provide strong evidence for the existence of voiced initials in the Pre-Miao-Yao language. For instance:

	I	III	V	VIII
	'weed'	'long'	'wear hat'	'scoop'
MKC	ntou ¹	ntæ ³	nton ⁵	ntše ⁵
MKS		nta ³	ntaə ⁵	ntšao ⁵
YLP		nta ³	ntou ⁵	
MTK		te ³	tau ⁵	tšu ⁷
MCF		ta ³	tau ⁵	tšu ⁷
MJC	ton ¹	tai ³	tu ⁵	tšo ⁷

	II	IV	VI	VIII
	'yam'	'leech'	'hemp'	'pluck fruit'
MKC			ntu ⁶	
MKS			nto ⁶	
YLP	ntau ²	ntsi ⁴	nti ⁶	nti ⁸
MTK	na ²	ñi ⁴	no ⁶	ñe ⁸
MCF	na ²			
MJC	na ²	ñi ⁴	nau ⁶	ñi ⁸

The difference in correspondence among words with the first and the second tone type suggests that the initials of the former were originally voiceless, while the initials of the latter were probably voiced (thus, **mb-* > **m-*, **nd-* > **n-*, **ndz-* > **n-*, **ndž-* > **ñ-*, **nd-* > **n-*, **ng-* > **ŋ-*, **ŃG-* > **Ń-*, etc.). This kind of reconstruction fits into the system of Pre-Miao-Yao initials as a whole.

Once voice has been reconstructed as a phonemic feature in the initials of words with the second tone type, this type becomes superfluous in the Pre-Miao-Yao pattern;³ the eight categories of tones are reduced to four. It is the voiced nature of the initials that caused the development of four new tones in a later stage of the language.

The unsolved problem is that of relating the odd-numbered tones I, III, V with the even-numbered tones II, IV, VI in order to establish three categories of tones for Pre-Miao-Yao. My tentative solution is suggested by the following three examples, which may reflect a historical relationship⁴ between I and II between III and IV, and between V and VI.

	I	II	III	IV	V	VI
		'fall'		'wear clothes'		'earth'
MKC	poŋ ¹		nhaŋ ³		loŋ ⁵	
MKS	pəə ¹		nhoŋ ³		laə ⁵	
YLP	pou ¹			nan ⁴	lou ⁵	
YYT		pa ²	ne ³			
MTK		pa ²		naŋ ⁴		la ⁶
MCF		pai ²		neh ⁴		la ⁶
MJC		pa ²		nəŋ ⁴		la ⁶

Reconstruction results in a triad system of **b-*, **p-*, **ph-*; **bl-*, **pl-*, **phl-*; **br-*, **pr-*, **phr-*; **d-*, **t-*, **th-*; **dl-*, **tl-*, **thl-*; etc. As I have indicated, Pre-Miao-Yao had three types of initials: voiced, voiceless, and aspirated voiceless, occurring in words with only four different tones. In a later stage of the language, the voiced initials became voiceless; their original voiced quality results in the rise of four new tones parallel to the four old ones. As a result, such an initial as *th-* (from original **th-*) appears today only in words of the first tone type, whereas *t-* (from original **t-* and **d-*) appears in words of both types.

Nasal and lateral initials are of three series: (A) *mh-*, *nh-*, *ñh-*, *lh-* in words of the first tone type; (B) *m-*, *n-*, *ñ-*, *l-* in words of the first tone type; and (C) again *m-*, *n-*, *ñ-*, *l-* in words of the second tone type. There are various possibilities for the prototypes of these three series; we may reconstruct either (A) **ŋh-*, **ŋh-*, **ñh-*, **th-*, (B) **ŋ-*, **ŋ-*, **ñ-*, **l-*, (C) **m-*, **n-*, **ñ-*, *l-*; or else (A) **mh-*, **nh-*, **ñh-*, **lh-*, (B) **m-*, **n-*, **ñ-*, **l-*, (C) **m-*, **n-*, **ñ-*, **l-*.

It is, then, my working hypothesis that the Pre-Miao-Yao language had initials of three types (voiced, voiceless, and aspirated) and four tones, which later

⁴ For instance, alternation between voiceless and voiced initials.

developed into initials of only two types (voiceless and aspirated) and eight tones. The change may be represented in a diagram:

ORIGINAL STATE		LATER STATE	
TONES	INITIALS	TONES	INITIALS
A, B, C, D	voiced	II, IV, VI, VIII	voiceless
A, B, C, D	voiceless	I, III, V, VII	voiceless
A, B, C, D	aspirated	I, III, V, VII	aspirated

POPULAR CHINESE PLANT WORDS
A DESCRIPTIVE LEXICO-GRAMMATICAL STUDY

YUEN REN CHAO

University of California, Berkeley

GENERAL

0.1. The present study is primarily a description of the everyday Chinese words which the writer uses in connection with plant life and which he believes to be representative of Standard Mandarin. No attempt has been made to contribute to new identification of Chinese plants, though note will be taken of cases of conflicting authorities. The words covered include not only specific plant names such as 'rose' and 'pine', but also such general words as 'tree' and 'grass' and names of plant parts such as 'root' and 'sap', as well as verbs for what plants do, such as 'to blossom', and what man does to plants, such as 'to plant'.

0.2. While there is always some correlation between things and words,¹ so that a plant name is likely to be a noun and the word for a change occurring in a plant is likely to be a verb, the parallelism is imperfect and the proper study of words is still words.

0.3. One thing that linguists emphasize, and practical language students should give more attention to than they usually do, is the principle of total accountability. It is not enough to record the root meaning of simple morphemes. Account must be taken of all its variant forms and the environments in which it occurs. Botanical and even general dictionaries may give *li* (see 0.10 for the transcription; references throughout are to sections of this paper) for 'pear', *shing* for 'apricot', and *lii* for 'plum', but these fruits are, in the actual language, called respectively *li* (morpheme word), *shienql* (root plus suffix, accompanied by automatic alternation of central vowel with original high vowel), and *lii.tzy* (root plus suffix *-tzy*). Nobody calls a plum **lii* or **lieel*, and the form *shing.tzy* 'apricot' exists only in the dialects. Again, *jen.tzy* 'hazelnut' takes the suffix *-tzy*, and *liantzzy* 'lotus seed' is a simple compound with *-tzyy* in nuclear or primary position, while *songtzeel* 'pine nut' is a compound with the second element *-tzyy* suffixed by the retroflex suffix *-l*, resulting in the mid-vowel form *tzeel*.

0.4. Compounds, moreover, are not made up for botanical reasons only, but have grown up by linguistic usage. Not all leaves that turn red are called *hornq-yeh* 'red leaf'; only the maple—the tree as well as the leaf—is so called. The plant sesame or its seed is called *jy.ma*, but the oil and jam made from the seed are called respectively *mayou* (or *shiangyou* 'fragrant oil') and *jy.ma-jianq*, not **jy.ma-you* and only rarely *majianq*. (The only thing commonly called *majianq* is the game of 'mahjong'.) The leaf of the mulberry is called *sangyeh* 'mulberry-leaf', a compound word, and the leaf of the lotus is called *heriyeh* 'lotus-leaf', a similar compound. In speaking of the leaves of most other plants, however, such compounds are not used. Thus the leaf of an elm tree is simply called *yushuh*

¹ See M. B. Emeneau, *Language and non-linguistic patterns*, *Lg.* 26.2.199-209 (1950).

de yeh.tzy 'elm-tree's leaf', and a form like *yuyeh* 'elm-leaf', if understood at all, would sound unnecessarily poetic. A collection of things is not always named by a collective noun, and vice versa. Thus *shulin.tzy* 'woods' takes the individualizing auxiliary noun (AN) *-geh*, while a plant (in general) is only expressed by the collective noun *hua-tsao-shuh₀muh* 'flowers, grass, trees' (see 2.4), unless one uses the learned word *jyr₀wuh*. While one says *yi-ball huaball* for 'a petal' (2.45.1), one does not often say *gelgel gel* for 'every root' (2.97)—at least not much oftener than one says in English 'A rose is a rose is a rose'.

0.5. Such facts are of no botanical significance, and a botanical dictionary is in no way incomplete for not recording them. Nor can philological dictionaries, concerned with the etymologies of single morphemes, be expected to contain plant words in the full spoken form. But somewhere the facts about what is done and what isn't done in talking about plants must be recorded; the present study is a first draft of such a record.

0.6. The source of information in this study is first of all the writer's own speaking-knowledge of the language. Since he has lived in many places where non-Mandarin dialects are spoken, some of his vocabulary must have been influenced by them. As far as he is aware, all such forms are marked 'dial.' or by the place names where the non-Mandarin forms are used. Some 'dial.' items probably escaped detection, but they cannot be numerous.

0.7. Although this is not a study of realia, the scope of the paper is limited by including only words for what the writer has actually seen: plants or parts of plants, such as certain herbs or lumber trees. Plant words which the writer has read or heard about, or only talked about, are not included. A number of plant words from books are included for comparison or for the sake of their interesting constructions. Such forms are put in parentheses in the lexicon. The advantage of this restriction lies in that it puts the writer also in the position of an anthropological informant.

0.8. ABBREVIATIONS:

adj., adjective
AN, auxiliary noun, classifier
B, bound morpheme
BB', compound or morphemes *B* and *B'*
bin., binom, two syllables which, at the descriptive level, form one morpheme
B-l, derived word with non-syllabic retroflex suffix
BN, see 3.0
B.tzy, derived word with syllabic suffix-.*tzy*
Ch, see 3.0
CHPT, see 3.0
collec. n., collective noun
dial., nonstandard, but occurring in the writer's speech
F, free morpheme, morpheme word
fl, flower
fr, fruit

identif., identification
i.e., the term placed before this abbreviation is a literal translation and the term after it is an equivalent translation
-l, the retroflex non-syllabic noun suffix with value [-*l*]
L, learned or literary form (\neq *L.* for 'Linnaeus', following species names)
Lf, see 3.0
n., noun
pl, plant
Rd, see 3.0
sd, seed
.tzy, noun suffix
Unid., unidentified (botanically)
v.i., intransitive verb
v.t., transitive verb
Wu, dialect region of Chekiang and southeastern Kiangsu

0.9. SYMBOLS: A dot before a syllable indicates that it is in the neutral tone, i.e. toneless; a subscript zero (₀) marks an optional neutral tone; a hyphen marks

a simple adjective-noun compound (not a noun-noun compound), and separates the immediate constituents of a complex compound; part of a form in parentheses is an optional element; longer material in parentheses consists of dialectal or learned forms or quoted forms for which the writer is not reporting as informant; square brackets enclose IPA symbols; a breve over a vowel (ă) indicates an 'Entering Tone' in a dialect word, which loses its glottal stop when in nonfinal position; an equal-sign means 'identified with' (used with English names); a mark of inequality (\neq) means 'not the same as' or 'to be distinguished from'; a curve (\sim) indicates alternation, usually between variant forms of a character; a vertical line denotes repetition of a character.

0.10. ROMANIZATION. Because of the great number of citations, National Romanization is used here for the sake of typographical simplicity. For a summary of the spelling system see Y. R. Chao and L. S. Yang, *Concise dictionary of spoken Chinese* xx-xxi (Cambridge, Mass., 1947). The principal features of this system are as follows:

- (1) Unaspirated voiceless stops are *b, d, g, t, k, p*.
- (2) Aspirated voiceless stops are *p', t', k', t', k'*.
- (3) Wade *-ien, -ih, -ü, -rh, -ung, -ü* are written respectively *-ian, -y, -y, -l, -ong, -iu*.
- (4) For 1st Tone (high-level) with initials *l, m, n, r*, write *lh-, mh-, nh-, rh-* respectively; for 1st Tone with other initials use the basic unmodified spelling.
- (5) For 2nd Tone (high-rising), add *-r* after a vowel or use *y* or *w* if the syllable has a medial, but use the basic form if the initial is *l, m, n, r*.
- (6) For 3rd Tone (low-dipping), double the vowel or change the combinations *ai, au, ia, io, iu, ua, ue* to *ae, ao, ea, eo, eu, oa, oe* respectively.
- (7) For 4th Tone (high-falling-to-low), change endings zero, *-i, -u, -ng, -l* to *-h, -y, -w, -ng, -ll* respectively.
- (8) A period BEFORE a syllable indicates that the syllable is unstressed.

1. THE STRUCTURE OF PLANT WORDS

1.0. The present section deals with the types of internal structure in plant words from a purely grammatical point of view. These types will in general cut across the aspects of plant life as realia, but not without some correlation of various degrees. Thus, the form *B-l* (a noun formed of a simple morpheme plus retroflex suffix) can almost always be used for plant parts, but never for whole plants, general or specific. Such facts will of course be noted.

1.1. MONOSYLLABIC MORPHEME WORDS (marked *F*) are rare. Of the words included here there are only 6 cases of specific plant names: *daw* 'rice', *jiang* 'ginger', *tsong* 'scallion', *suann* 'garlic', *shen* 'ginseng', *ay* 'moxa'; 13 words for specific plant parts and plant products, e.g. *mii* '(hulled) rice', *soen* 'bamboo shoot', *oou* 'lotus stem', *chi* 'lacquer'; 5 general plant names: *tsay* 'vegetable', *tsao* 'grass', *yaw* 'medicine, herb', *ternng* (*F* dial.) 'vine', *shuh* 'tree'; and 12 general plant parts, e.g. *gen* 'root', *shiu* 'hair, awn', *pyi* 'skin', *you* 'oil'. Note that the proportion of general to specific words is almost half and half; this is a

much larger percentage for general words than in the category of compounds, where specific names predominate.²

Verbs about plants (2.106) are mostly monosyllabic, but they are usually tied up with some aspect suffix in actual use. Adjectives (2.123) are also mostly monosyllabic, though few cited here are restricted to application to plants as are some verbs and most of the nouns.

1.2. BINOMS. At the descriptive level, which is the level of this study, a binom is an unanalyzed dissyllable, whether it was once analyzed or not. Conversely an original binom may have acquired separate meanings and then should be regarded as a compound, even though it is one only through 'folk etymology'.

Binoms figure very largely in dissyllables of the trochaic type, i.e. those with full stress on the first syllable and neutral tone (zero stress) on the second. There are far fewer binoms, both relatively and absolutely, in dissyllables of the quasi-iambic type, i.e. those with half stress on the first syllable and full stress on the second. Thus, while binoms form only 4% of quasi-iambics, they form about 25% of the trochees.

1.3. DERIVED WORDS. Almost all derived words about plants are formed by suffixation: there are no prefixes and only a couple of doubtful cases of reduplication. Color words, taste words, etc., which are often said to be 'prefixed' to plant words, are really first members of compounds and not true prefixes. The two commonest suffixes in plant words are *-l* and *-tzy*, both coming from words meaning 'child'.

1.3.1. Of the 25 derived words with suffix *-l* there are 3 words for fruits, e.g. *taurl* 'peach'; 19 (incl. *gu.dul* '(flower) bud', which is a suffixed binom) are words for other plant parts, e.g. *hual* 'flower'; and 3 are abstract nouns from AN's (see 2.69). The proportion of general to special words is about $\frac{3}{4}$. The remarkable thing is that except for plants named by their flowers or fruits—as *Jeh shyh yih-ke taurl(-shuh)* 'This is a peach (tree)'—not a single word for the whole plant, general or specific, takes the form of *B-l*.

1.3.2. The suffix *-tzy* is to be distinguished from *-tzyy* 'seed', or its *l*-suffixed form *tzeel* 'seed', is in *liantzyy* 'lotus seed' and *quatzeel* 'watermelon seed'. The suffix occurs in plant words of the following categories: (a) grains, as *yih-ke may.tzy* 'a wheat plant', *yi-liell may.tzy* 'a grain of wheat'; (b) fruits and nuts, as *yih-ke lii.tzy* 'a plum (tree)', *yi.geh lii.tzy* 'a plum'; (c) other plants, as *jwu.tzy* 'bamboo'. Note that a tree bearing a fruit whose name has the suffix *-tzy* is normally called such-and-such *shuh* 'tree', in preference to the fruit name alone; but grains and other plants have no names except the forms with *-tzy*, nothing else being omitted or 'understood' after *may.tzy* 'wheat' or *jwu.tzy* 'bamboo'.

1.3.3. A few plant parts are named with *-tzy*. The ratio of general to specific words with *-tzy* is about $\frac{1}{3}$, as compared with $\frac{3}{4}$ for words with the retroflex suffix *-l*.

1.4. DE-COMPOUNDS. A de-compound is a word consisting of a compound which, by virtue of taking on an affix, ceases to be a compound and becomes a derived

² On the cultural primacy of morpheme words for plants see Lf 547, and for things in general see *Selected writings of Edward Sapir* 92-3 (ed. D. G. Mandelbaum; Berkeley, 1949).

word. A Chinese example is *paotarngl* 'runhall-er, one who runs around the dining hall, i.e. a waiter'. It is not a compound *pao-tarngl*, since there is no such word as *tarngl*. The suffix *-l* 'er, one who' is added to the verb-object expression *pao-tarn* to make of it an agent word. Among plant words, however, there seem to be no de-compounds. All forms of the type *BB'-l* seem to resolve into true compounds of *B* and *B'-l*. (Cf. 1.12 on exocentric compounds.)

1.5. SIMPLE COMPOUNDS. We have seen that a large proportion of free morphemes and derived (suffixed) words are general words for plants and plant parts. Compounds, on the other hand, are the typical form for specific plant names. The commonest form of compound has a first element in subordinate position to the second element; it is in the form either of noun-noun, as *woasong* 'tile pine, i.e. roof pine', or of adjective-noun, as *shiang-tsay* 'fragrant vegetable, i.e. coriander'. A large proportion of the noun-noun compounds take the form of a specific name followed by the general name in close apposition, as *songshuh* 'pine tree'. This is the same kind of compounded or close apposition as in '*dah*'-*stzyh* 'the *dah*-character, i.e. the character *dah*', or *Wang.shian.sheng* 'the Wang gentleman, Mr. Wang'.

1.6. The preceding types cover the majority of simple compounds. Of the approximately 200 forms recorded here, some 80, or 40 %, are noun-noun compounds; 26, or 13 %, are compounds by close apposition; 54, or 27 %, are adjective-noun compounds. Of the remaining 40 forms or 20 %, the most interesting are exocentric constructions, such as *farn-g-feng* 'ward-wind', which is a verb-object compound, used substantively as if standing for *farn-g-feng .de yaw* 'ward-wind herb'. If we regard the unexpressed substantive as Jespersen's zero,³ so that the construction is analyzed as 2 (V O) 1⁰, this may be regarded as a complex compound with a zero second member. We can similarly classify *ying-chuen* 'welcome the spring' (No. 160), which as a matter of fact does alternate with *yingchuen-hual* 'welcome-the-spring flower'. A very peculiar formation is *huasheng* 'peanut' (No. 210), which is an abbreviation of the less common form *lawhua-sheng* '(when) fall the flowers (the nuts) grow', because the nuts grow from fallen flowers.

1.7. The most frequent first element entering into compounds denote the following categories:

1.7.1. Colors: *horng* 'red', *hwang* 'yellow', *ching* 'green',⁴ *tzyy* 'purple', *bair* 'white', *hei* 'black'. Note the comparative infrequency of *lan* 'blue' and *liuh* 'green'.

1.7.2. Other qualities: *dah* 'large', *sheau* 'small', *shiang* 'fragrant', *chow* 'stinking', *tyan* 'sweet', *suan* 'sour', *kru* 'bitter', *lah* 'hot, pungent', *tsao* 'grass-', *muh* 'tree-'.

1.7.3. Place and time of growth: *shan* 'mountain', *shoei* 'water', *yee* 'wild', *yang* 'foreign', *shi* 'west', *nan* 'south' (but only rarely *dong* 'east' and *bee* 'north'), *chuen* 'spring', *shiah* 'summer' (rare), *chiou* 'autumn', *dong* 'winter'.

1.8. The commonest second elements are: *-dow* 'bean', *-tsay* 'vegetable', *-gua*

³ Otto Jespersen, *Analytic syntax* 16-7 (Copenhagen, 1937).

⁴ This is its usual meaning as a combining form in plant words. As a free word, *ching* means 'blue-green, light blue'.

(usually in the form *-gua* 'melon', *-tsao* 'grass', *-hua* 'flower' (but see 1.9.2), *-guoo* 'fruit', *-tour* 'head, tip', *-shuh* 'tree', *-muh* 'tree', *-tzyy* 'seed'. (Cf. also 1.9.3.)

Before leaving the subject of simple compounds let it be noted that among plant words, unlike the general vocabulary, coordinate compounds, whether of parallel items or opposites, are for some reason very rare. Examples are: *guaguoo* 'melons and fruits' (collec. n.), *luwoei* 'reed-reed, i.e. common reed', *hua-tsao-shuh, muh* 'plants' (collec. n.).

1.9. COMPOUNDS WITH SUFFIXED ELEMENTS. One or each element of a compound may itself be a derived word with a suffix. Compounds in which both elements have suffixes are rare. There are *jiy.tzyhual* 'gardenia', *banq.tzy-miall* 'corn meal', and if reduplication is regarded as affixation, then *gar.ga-tzaol* 'jujube (with a rhombic instead of the usual oval longitudinal section)'.

1.9.1. Compounds with a suffixed first element are also relatively rare, as *pyaul-tsay* (No. 39), *jiaul-tsay* (No. 48), *ial-li* 'Chinese pear'.

1.9.2. This leaves us with the suffixed second element as the commonest type under this head. The suffix may be *-tzy* or *-l*, and the second syllable may or may not be stressed (resulting respectively in quasi-iambic or in trochaic compounds). The four possibilities and the number of occurrences are as follows:

<i>-B.tzy</i>	6 cases	example <i>shuh-yeh.tzy</i> 'tree-leaf'
<i>-B.tzy</i>	none	
<i>-B-l</i>	44 cases	example <i>jiuhual</i> 'chrysanthemum (fl)'
<i>-B-l</i>	2 cases	example <i>yu.chyal</i> 'elm seed'

The word *gu.dul* '(flower) bud' looks like a compound with an unstressed suffixed second element; but since it is a binom, it is to be regarded as a derived primary word and not a compound, except in the sinological sense of a compounded group of characters. (As a matter of fact this is one of those very common spoken words for which there is no settled way of writing them in characters.)

1.9.3. Note that the simple second-element forms *-dow* 'bean', *-hua* 'flower', and *-guoo* 'fruit' (listed under 1.8) are much less frequent than the corresponding suffixed forms *-dowl*, *-hual*, and *-guool*. The form *-guool*, moreover, has a wider application in that it also includes 'berries' and 'nuts'.

1.10. COMPLEX COMPOUNDS. Most of the complex compounds among the words recorded here have immediate constituents of which only one is a compound. The following general features can be noted.

1.10.1. Among compounds of three morphemes with no neutral tone, there is a preponderance of two-plus-one over one-plus-two forms. Thus, there are about 10 cases like *yang-bairtsay* 'foreign-celerycabbage, i.e. cabbage', as against some 50 cases like *maawoei-song* 'horsetail pine'.

1.10.2. If there is a neutral tone, it serves as a strong binder for the subcompound, so that *B.B'B"* is always *B.B'-B"*, while *BB'.B"* is always *B-B'.B"*, for example *jiy.jia-tsao* 'fingernail grass, i.e. balsam', but *yee-her.taur* 'wild nut-peach, i.e. wild walnut, i.e. hickory'. Apart from exocentric constructions (see 1.12) there seem to be no exceptions to this rule.

1.10.3. If the last element has a suffix, then, irrespective of the occurrence or position of a neutral-tone syllable, the construction is almost always two-plus-one and not the other way, for example *kweihua-tzeel* 'sunflower-seed', *liuh.dow-*

yal 'green-bean sprout, i.e. mungbean sprout' (the sprout being actually white). An apparent exception is *hua-gu.dul* 'flower bud'; but *gu.dul*, as we have seen, is not a compound, but only a suffixed dissyllabic morpheme, and so the compound *hua-gu.dul* is not complex. A borderland case is *bair-lanhual* 'white lan-flower', a Wu dial. name for a regional flower little spoken of in the north (see No. 158).

1.11. Auxiliary nouns (AN) and verbs will be treated in sections 2.90-137, where detailed lists are given.

1.12. EXOCENTRIC CONSTRUCTIONS. A look through a compilation of plant names such as Ch reveals a great many in the form of predicates, or of full sentences with subjects and predicates. Used as plant names they are of course exocentric constructions, that is, constructions in which there is no nucleus which has the same grammatical function as the whole. Such names are actually not so common in popular usage as the encyclopedias would indicate. In 1.6 we have noted predicates like *ying-chuen* 'welcome spring', with a following zero substantive. Examples of longer predicatives are *say-sheuei* 'excels snowpear' (No. 51.3), name of a large green radish eaten as fruit (cf. 1.12.2); *yueh₂yuell-horng* 'monthly red, monthly blooming rose' (147); *chuen.bulao* '(in) spring not tough' (29), one of the rare cases where a neutral tone does not bind its syllable closer to the preceding; *wannnian-ching* '10,000-year green' (137).

1.12.1. The form *dong-chorng-shiah-tsao* 'winter-insect-summer-grass' (101) should not be regarded as an attributive compound, but rather as two predicates 'winters (it's) insect, summers (it's) grass'. As a matter of fact there is an alternate form with *-jiunn*, here in the sense of 'plant', as the nucleus.

1.12.2. Names consisting of full sentences describing the nature or appearance of plants are not common in colloquial usage, as we have noted. Examples from Ch are: *tiee-shuh-kai-hua* 'iron-tree-opens-flowers' Ch 204, *jyju-baw-dann* 'spider-embraces-eggs' Ch 214, *shian-ren-guoh-chyau* 'immortals-cross-the-bridge' Ch 215, *wang-buh-liou-shyng* 'king-does-not-keep-(you)-from-going' Ch 271 (No. 140), none of which the writer is acquainted with. The only live example in the present lexicon is No. 51.3 (191.2 being ambiguous in construction) *luo.bo-say-sheuei* 'radish-excels-snowpear', which is an alternate form to the one cited in 1.12 above. Since there seems to be a nucleus *luo.bo* 'radish', that would make it endocentric again. But then to place an attributive element *say-sheuei* after the nucleus would be contrary to the fundamental principle of Chinese construction. It is therefore still preferable to regard this as an exocentric construction.

2. GENERAL PLANT WORDS

2.1. CLASS NAMES. Under this heading are included words like *tsao* 'grass', *shuh* 'tree', which are of course not necessarily classes in the botanical sense. There is no popular word to cover all plants. The term *jy₂wuh*, though often used when one speaks with conscious generality, still sounds classroomish and is therefore marked L in the following list. There is however a general AN *-ke* for words of any degree of generality so long as they refer to individual plants, as *yih-ke tsao* 'a grass,' *yih-ke song* 'a pine.' The AN for a 'kind' is *-joong* (again not in the botanical sense of 'species'), as *yih-joong tsay* 'a kind of vegetable'.

Since these two AN's are applicable to all words for plants, they will not be listed under each separate plant name, except as a reminder when a more specific AN is given as an alternate form.

(2.2. *gyrawuh* 'plant' *L.*) In compounds such-and-such '-plant' is called *-tsao*, e.g. in No. 133.

(2.3. *tsaomuh* 'grasses and trees, plants' *collec. n., L.*)

(2.3.1. *yih-tsao yi-muh* 'a single plant' *L.*)

2.4. *hua-tsao-shuh-muh* 'flowers, grasses, trees, i.e. plants' *collec. n.*

2.4.1. *muhbeen.de* 'of tree root, of woody root' largely overlapping 'perennial'.

2.4.2. *tsaobeen.de* 'of grassy root' largely overlapping 'annual'.

2.5. *guu.tzy* 'grains' pl or sd.

(2.6. *wuu-guu* 'the five grains, the grains' *collec. n., L.* The list of five grains has varied from age to age.)

2.7. *tzar-llang* 'miscellaneous grains' *collec. n.*, term applied to grains other than wheat and rice.

2.8. *dowl, dow.tzy* 'bean' pl or sd.

2.9. *tsay F* 'vegetable.' (The word *tsay F* is also used in the sense of 'cooked dish', whether vegetable or animal.)

2.10. *huen-tsay*. (Term applied to vegetable with strong smells, such as garlic, onion, but usually not to spices. But a more frequent use of the term is for cooked dishes containing animal food. *Huen-tsay* of both kinds are tabooed by religious vegetarians.)

2.11. *wuu-huen, collec. n.* for No. 2.10, first meaning. Authorities differ as to what five are included in the list of five.

2.12. *shiang-liaw* 'spice', word applied only to the material and not to the plant. It is applied to the seed in so far as the spice is used in the form of whole seeds, as *hwujiau* '(black) pepper'. The compounded form *wuu-shiang* 'five-spice, all-spice' is always bound. The complete adjective word is *wuu-shiang.de*.

2.13. *suh-tsay* 'vegetarian vegetable, i.e. vegetable'. Also applied to cooked dishes containing no animal food. The term does not cover staple food from plants used for its starch, such as rice and wheat.

(2.14. *tsay-su*, dial. and *L, sutsay L* 'vegetables' *collec. n.*)

2.15. *tsao F* 'grass' AN for 'blade' *-gel*. In compounds, either in first or in second position, it is used in contrast with *shuh* or *muh* 'tree'.

2.15.1. *ching-tsao* 'green grass, i.e. grass', term applied to any kind of green grass. On the other hand *liuh-tsao* 'green grass' is not the usual term for grass, but a made-up slightly picturesque expression. (Cf. 1.7.1.)

2.16. *yaw F* 'medicine'. As a morpheme word, *yaw* refers more often to the prepared form or part of a medicinal plant than to the plant. Thus the sentence *Day.hwang shyh yih-joong yaw* is more likely to be understood as '(The prepared form of) rhubarb is a medicine' than as 'Rhubarb is a medicinal plant'.

2.16.1. *yawtsao* 'medicinal plant'.

2.17. *terng F* (dial.) 'climbing plant, vine' AN *-joong*.

2.17.1. *terng.tzy* 'climbing plant, vine' AN *-ke, -joong*.

2.18. *hual* 'flower' fl or pl, AN for fl *-duoo(l)*, AN for pl *-ke*, e.g. *jonq yih-ke hual* 'to plant a flower (plant).'

2.19. *shoeiguoo(l)* 'water-fruit, i.e. fruit,' *shian-guoo(l)* 'fresh fruit, i.e. fruit' AN for fruit *-geh*, AN for kind *-joong*. Note that although the name of a specific fruit can be used also for the plant, as *yih-ke shienql* 'an apricot (tree)', one does not usually speak of *yih-ke shoeiguoo* for 'a fruit tree'. (Cf. 2.20.) Note also that while the AN for the general word *shoeiguoo* etc. is *-geh*, the AN for words for specific fruits may sometimes take other forms. Thus, *yi.geh li* 'a pear', but *yih-gel shiang-jiau* 'a banana'. This is just another instance of the fact that AN's, like genders, have to do with words and not with things.

2.20. (shoe) *guoo-shuh* 'fruit tree'.

(2.21. *guannmuh* 'shrub' *L.* There is no popular word for shrubs, except such circumlocu-

tions as *ae-shuh* 'short tree', *sheau-shuh* 'small tree', *yih-tswan sheau-shuh* 'a bunch of small trees, i.e. a bush'.)

2.22. *shuh* *F* 'tree'.

2.22.1 *shuh.muh* 'trees' *collec. n.*

2.22.2 *shuhlin* (dial.), *lin.tzy*, *shuhlin. tzy* 'woods' *AN* *-geh*, *-piann*.

2.22.3. *s(h)enlin* 'forest' *AN* *-piann*, *-geh*.

2.22.4. *-muh* 'wood, -tree' in names of utility trees, such as Nos. 260 and 264.

2.23. PARTS OF PLANTS. The words for parts of plants have proportionately fewer compounds than class words or plant names. They are like names of other objects in being mostly either morpheme words, such as *tsyh F* 'thorn, prickly', or primary derived words (bound morpheme with suffix), such as *yeh.tzy* 'leaf'. Binomial morphemes are rare. In the following list general words applicable to more than one kind of plant are given first, followed by names of parts peculiar to certain plants only.

2.24. *ke'l* 'size, shape, or condition of growth of a plant'.

2.25. *gen*, *gen.tzy* 'root' *AN* *-geh*, rarely *-gen*, *-gel* (Cf. 0.4). The word is applied to any underground or underwater part of a plant if it is popularly regarded as the root.

2.26. *shiu F* 'hair, awn' *AN* *-gel*.

2.27. *juang(.tzy)*, *shuhjuang(.tzy)* 'stump' *AN* *-geh*, *-gen*.

2.28. *-tour B* 'head,' as *tsaytour* 'enlarged edible part of a vegetable', *suanntour* 'garlic (bulb)', *tsongtour* 'head of scallion'.

(2.29. *gann L*, *shuhgann L* 'trunk'.)

2.30. *shen.tzy*, *shuh.de shen.tzy* '(tree's) body, i.e. trunk' *AN* *-geh*.

2.31. *muh.tou* 'wood' *AN* for piece *-kuay*, *AN* for log *-gen*, *-gel*.

(2.32. *jing*, *shyng* 'stem' *L*.)

2.33. *geeng.tzy*, *geengl* 'stem' *AN* *-geh*, *-gel*.

2.34. *gaan.tzy*, *gaal* 'stalk' (of grains), *AN* *-gen*, *-gel*.

2.35. *ing.tzy* 'tops', as *luo.bo-ing.tzy* 'radish tops'.

2.36. Cf. *-gan* in *chyi.gan* 'flagpole', *wei.gan* 'mast', *-gal* in *lan.gal* 'railing'; *gaal*, *gaan.tzy* 'handle, stem (of a made object)'; *gan.tzy* 'pole', *jwugal* 'bamboo pole', *jianngaal* 'shaft (of an arrow)'. (See also the characters associated with these words in Section 4.)

2.37. *terng(.tzy)* *terng F* (dial.), 'vine' (as part of pl, not including the leaves, etc., cf. 2.17) *AN* *-gen*, *-gel*.

2.37.1. *terngsy*, *terngsel* 'tendrill' *AN* *-gel*.

2.38.1. *shuhjy*, *jy.tzy*, *shuhjy.tzy* 'bough, branch' *AN* *-gen*, *-gel*.

2.38.2. *jy.tzy*, *shuhjel*, less often *jel* 'twig' *AN* *-gel*, *-gen*.

(2.38.3. *iacha*, *shuh-iacha* 'bough, branch' dial.)

(2.38.4. *iajy*, *shuh-iajy* 'bough, branch' dial.)

2.39. *yeh.tzy* 'leaf' *AN* *-geh*, *-jang* (for large ones).

2.39.1. *shuhyeh.tzy*, *shuhyell* 'leaf of a tree'.

2.39.2. *yell* 'size, shape, or condition of growth of leaves, foliage'.

2.40. *shuhpyl* 'bark' *AN* *-kuay*, *-jang*.

2.41. *tsyh F* 'thorn, prickly, spike' *AN* *-gel*.

2.42. *gu.dul*, *huagu.dul* 'bud (of a flower)' *AN* *-geh*.

(2.42.1. *bau.tzy* 'bud' dial.)

(2.42.2. *roei* 'bud' *L*. The Wu dial. form [nydɔw], sometimes interpreted as *neutour* 'female head', is of course just *roeiour*, since *roei* < Ancient Chinese *ńzwi* 'bud' would regularly be cognate with Wu [ny].)

2.43. *hual* 'flower' *AN* *-duoo*, *-joong*, *-ke*. Note that although the *AN* for words for specific flowers is either *-duoo* or *-duool*, that for the word 'flower' itself is *-duoo*, which is a result of the dissimilatory tendency not to have retroflex endings in successive words.

2.44. *huashiel* 'pistil and stamina (as a group)' *AN* *-tswan*.

- 2.44.1. *huashiu* 'stamen or pistil (individually)' *AN -gel*.
 (2.44.2. *shyongroei*, *shyong-huashiu* 'stamen' *L.*)
 (2.44.3. *tsyrroei*, *tsyr-huashiu* 'pistil' *L.*)
 2.45. ball 'size, shape, or nature of petals (less often of sections)'. (See also 2.102)
 2.45.1. *huaball* 'petal' *AN -geh, -ball*, e.g. *yi-ball huaball* 'a petal'.
 2.45.2. *danbann.de*, *danball.de* 'single (layer)-petaled'.
 2.45.3. *shuangbann.de*, *shuangball.de* 'double(layer)-petaled'.
 2.46. *huatuol* 'calyx'.
 (2.46.1. *eh* 'calyx' *L.*)
 2.47. *huafeen* 'pollen'.
 2.48. *guool*, *guoo.tzy* 'fruit' (i.e. relative to the pl. cf. 2.19 *shoeiguoo*).
 (2.48.1. *guoo*, *shyr*, *guooshyr* 'fruit (in the above sense)' *L.*)
 2.49. *rangl* 'pulp, meaty part of fr'.
 (2.49.1. *row(.tzy)* 'meat or quality (of fruit or nut)' dial.)
 2.50. *pyl F* 'skin, rind'.
 2.50.1. *pyel* '(thin) skin' (e.g. inner skin of peanut); 'thickness or quality of skin'.
 2.51. *gua F* 'melon'.
 2.52. *guaguoo* 'melons and fruits' *collec. n.*
 2.53. *dih. tzy* 'stem (of fr)'.
 2.54. ball (< *bah* + *-l*) 'stem (of fr or fl)'. \neq 2.45 *-ball* < *-bann* + *-l*.
 2.55. *-gal* 'dried (fr)', as *pwu.taur-gal* 'raisin'.
 2.56. *hwul* 'stone'.
 2.57. *tzeel* 'seed' *AN -geh, -lih, -ke* (\neq 2.931).
 2.58. *joong F* 'seed, cion' (for planting). (See also 2.92.)
 2.58.1. *joong.tzy* 'seed'.
 2.59. *ker'l* 'shell'.
 2.60. *rel* 'kernel'.
 2.61. *guu.tzy* '(unhulled) grain' (also applied to the pl, cf. 2.5).
 2.62. *suell* 'tassel, ear (of grains)'.
 2.63. *dowl* 'bean' *AN -lih*.
 2.64. *dowjeaul* 'pod'.
 (2.64.1. *jya* 'pod' *L.*)
 2.65. *yal* 'bud, sprout'.
 2.65.1. *miaul* 'shoot, sprout'.
 2.65.2. *lang F* 'shoot, seedling, cion'.
 (2.65.3. *-tourl* 'shoot, sprout' dial., as *uandow-tourl* 'green shoots from peas' = *uan.dow-miaul* No.19.)
 2.66. *jiang F* 'sap'.
 2.67. *jy F*, *jel* (dial.), *jy.tzy*, 'juice'(from any part of the pl).
 2.68. *you F* 'oil'.

2.69. ABSTRACT NOUNS. As can be seen from the preceding entries, a suffix *-l* (less often *-tzy*) added to the name of a plant part makes of it an abstract noun meaning 'size, shape, quality, etc. of the plant part'. To resume, we have *ke'l* for the pl, *gaangl* for the stem, *jel* or *jy.tzy* for branches, *yell* (less often *yeh.tzy*) for foliage, *ball* for petals, *hual* for fl, also *huaduool*, which, in addition to the usual abstract meanings, also refers to the grouping of blossoms.

2.70. SPECIAL LIST. The items 2.71 to 2.90 consist of words peculiar to the parts of certain plants only (except of course names of flowers and fruits, many of which stand for the plant names for short). For details on special products from plants, such as soybean sauce (15) and reed matting (123), see under the specific plants, Section 3.

- 2.71. mayya(l) 'malt'.
 2.72. kang, longkang 'husk, chaff (of wheat or rice)'.
 2.73. fu.tzy, fu.pyi, mayfu(.tzy) 'bran'.
 2.74. mli (hulled, but uncooked) rice'.
 2.74.1. milliell 'grain of rice'.
 2.75. shwu.ji ~ shwu.jie 'edible sweet stalk of sorghum'.
 2.76. (liuh.)dowyal, chiatsay 'peasprout, sprout of mungbean'.
 2.77. dowyal(tsay) 'soybean sprout'.
 2.77.1. dow.fu 'bean curd'.
 2.77.2. dow(.fu)-jlang 'soybean milk'.
 2.78. tsaytair 'tips of vegetable bearing flowers or buds' (esp. of *youtsay* No. 25.).
 2.79. seen *F* 'bamboo shoots', in combining form also applied to shoots of other plants, as *lusoan* 'asparagus (tips)'.
 2.79.1. jwum(~n)leh(.tzy) 'bamboo splints' *AN -gel, -tyau, -piall*.
 2.80. oou *F* 'edible rootstock of lotus' *AN -gel*.
 2.80.1. lian.perng '(whole) fruit of lotus'.
 2.80.2. liantzyy '(fresh or dried) lotus seed', lian.perng-tzeel '(fresh) lotus seed'.
 2.80.3. heryeh 'lotus leaf'.
 2.81. guatzeel 'melon seed', usually applied to that of watermelon.
 2.82. tarn *F* '(cane) sugar'.
 2.83. songjen, songjel 'pine needle' *AN -gel*.
 2.83.1. song.shiang 'resin'.
 2.83.2. songguool, songtaa 'pine cone'.
 2.84. yun.shiang 'essence of rue'.
 2.85. leoutyaul 'willow twig' (esp. barked twig for weaving baskets).
 2.86. yu.chyal 'elm seed' (so called because of a slight resemblance to a coin).
 2.87. sangrell 'mulberry (the berry)'.
 (2.87.1. sangshenn *L*.)
 2.88. chl *F* 'lacquer'.
 2.89. shuhjiau 'tree-glue, i.e. rubber, raw or prepared'.
 2.89.1. shianqpyi '(prepared) rubber'.
 2.90a. feng(shuh)tarn 'maple-sugar'. (Both the object and the compound word seems to be a recent borrowing.)

2.90. AUXILIARY NOUNS (AN). Auxiliary nouns, as the term suggests, are auxiliary to nouns. Thus, *shuh* 'tree' is a noun and takes the *AN -ke*, as *san-ke shuh* 'three trees'. An *AN* is not a free word, but always bound to a preceding demonstrative or numeral or to a suffix (2.69) to form a word. Some things about plants are not expressed by nouns but only by *AN*'s. Thus, 'kind, species (popular sense)' is expressed by the *AN -joong* (the *F* form under 2.58 is the word for the physical cion and not for the class), and 'bush, bunch' is expressed by the *AN -tsuan*. In certain contexts a noun can be used as a temporary measure of things and thus functions as an *AN*, e.g. *yi-shuh .de lihual* 'a treeful of pear blossoms', as against *yih-ke lishuh* 'a pear tree'. Such cases need not of course be listed.

2.91. *-geh*. This most general *AN* for words denoting individual entities is to be understood to be applicable to our lists unless some other *AN* is given. In general, nouns with a suffix *-l* or *-tzy* take *-geh* for an *AN*. Words for non-elongated fruits also take *-geh* for *AN*.

2.92. *-joong* 'species, kind' (not limited to the biological sense of *joong* 'species' *L*). In this wider, popular sense of 'kind', there is an alternate form *-tzoong*, probably a blend of *-joong* 'kind' and *-tzong* 'category, batch'. For simplicity we have used only the form *-joong* throughout, understanding that it often alternates with *-tzoong*. (This alternation is by no means limited to plant words.)

2.93. *ke AN* for words for any individual plant, to be distinguished from 2.94.

2.93.1. -jiah AN for vine on frame, e.g. *yi-jiah pwu.taur-tern* 'a grape vine (on frame)'.
(2.94. *ke* 'family' (in the biological sense) L.)

2.95. -jy 'branch, twig', as *yih-jy hual* 'a twig bearing (one or more) flowers'. This is not quite a temporary AN, since -jy is a bound form (cf. 2.38).

2.96. -jyel 'section', as *yih-jyel ju.tzy* 'a section of bamboo'.

2.97. -gel, -gen AN for stick-like parts of plants (or other objects), as *yih-gel may-gaan.tzy* 'a wheat stalk'. Note that although the root meaning of the AN -gen is 'root', the AN for the word for 'root' *gen F* or *gel* is -geh.

2.98. -jang 'sheet', as *yih-jang heryeh* 'a lotus leaf',

2.99. -duoo(l) AN for flowers, as *yih-duoo hwang-hual* 'a yellow flower', *yih-duool meigwey* 'a rose'.

2.100. -lih, -liell 'grain', as *yi-lih tsarndowl* 'one fava bean'.

2.101. -ke (a different morpheme from 2.93, see characters) 'grain', as *yih-ke miiliell* 'a grain of rice'.

2.102. -yal, -ball AN for natural sections of fruits or bulbs, as *yih-yal jyu.tzy* 'a section of tangerine', *yi-ball bae.her* 'a section of lily root'.

(2.103. -nang, same as 2.102, dial.)

2.104. -tswan 'bush, bunch', as *yih-tswan ae-shuh* 'a bunch of shrubs'.

2.105. -du.lu 'bunch', as *yih-du.lu pwu.taur* 'a bunch of grapes'.

2.106. VERBS. With all the wealth of species and varieties of plants, there are relatively few verbs specially associated with plants. Perhaps a great many statements about plants have taken the form of Aristotelian logic by using the colorless verb *shyh* 'is', while all the richness of predication has gone into the predicative noun, e.g. *Jeh shyh yih-ke meigwey* 'This is a rose (bush)', *Meigwey shyh yih-joong hual* 'The rose is a (kind of) flower'. The more specific kinds of verbal predicates are given below in the following order: (a) 2.108-122 What the plant or plant part does, (b) 2.123-126 The state or quality of the plant or plant part, (c) 2.127-137 What human beings do to the plant or plant parts. Since the main thing a plant does is to grow and decay, and since growth and decay are typical processes of change, the new-situation particle *-le* occurs with very great frequency, e.g. *Hual nhian .le* 'The flower has faded'. The denial of a *-le*-form is of course made by adding *mei* or *mei.yeou* and dropping the *-le*, as *Hual mei nhian* 'The flower has not faded.'

2.107. When something occurs in a part of a plant, the verb denoting that happening is sometimes predicated of that part, or of the plant. In the latter case the word for the part is in the object position. For example, *Jeh shuh jaang yeh.tzy .le* 'This tree is growing leaves', or *Jeh shuh .de yeh.tzy jaang.chu.lai .le* 'The leaves of this tree are coming out'. Where both forms are possible the first form, the verb-object form, is usually preferred.⁵ For simplicity only one translation will be given where two forms are possible in the Chinese.

2.108. *hwo .le* 'lives' (after planting or after injury).

2.109. *jaang .le* 'has grown' *v.i.*

2.109.1. *jaang* 'grow' *v.t.*, as *Renshen jaang yeh.tzy .bu .jaang?* 'Does ginseng grow leaves?'

2.110. *chu-yal* 'to issue sprouts'.

2.110.1. *jaang-miaul* 'to grow sprouts'.

2.111. *guann-jiang .le* 'sends up sap'.

⁵ Cf. more general cases of such constructions as *Ta syi.le yi.geh perng.yeou* 'He lost (through death) a friend', *Ta perng.yeou syi.le* 'His friend died'. On this construction see Joseph Mullie, *The structural principles of the Chinese language* 1.160 ff. (Eng. tr. by A. C. Versichel; Peiping, 1932).

2.112. *fa-ching* 'to turn green'.

2.113. *fa-yal* 'to issue buds'.

2.114. *jaang yeh.tzy, yeh.tzy jaang .le* 'to grow leaves'.

2.115. *jaang gu.dul, gu.dul jaang .le* 'to grow (flower) buds'.

2.116. *kai-hual, hual kai .le* 'to blossom'.

2.117. *jie* 'to bear (fruit)'. The expression *jie-guoo.tzy* is slightly *L*. The usual object to this verb is the name of some specific fruit, as *Jeh shuh jie shyrliou .bu jie?* 'Does this tree bear pomegranates?' Note that *jyeguoo* 'result, effect' is a noun.

2.117.1. *jie tzeel* 'to bear seeds'.

2.118. *sheng* 'unripe', as *Gua tay sheng* 'The melon is too unripe'.

2.119. *shour .le* 'ripe(ned)', as *Rangl shour .le .mei.yeou? Shour .le.* 'Is the pulp ripe now? Yes, it is'.

2.120. *nhian .le* 'to fade'; 'to wither'.

2.120.1. *bay .le, shieh .le* 'to fall apart (of a flower or its petals)'.

2.120.2. *law .le* 'has fallen'.

2.120.3. *law yeh.tzy, yeh.tzy law .le* 'The leaves are falling, or have fallen'. Note that the form *luoh-yeh* 'fallen leaves' is *L*.

2.121. *ku .le* 'to become dried up'.

2.122. *syy .le* 'to die'.

2.123. STATE AND QUALITY. Since state and quality are expressed by adjectives and since Chinese adjectives are verbs, any such word can be predicated of plant words. Thus *Ta mae .de tsay nenn* 'The vegetable he bought was tender'.

2.124. Adjectives can be made a complement ('co-verb') after the verb *fa* 'develop', with the resulting meaning 'to turn or give the impression of being . . .', as *fa-bair* 'to turn white or appear whitish', usually in a neutral or bad sense. Colors and tastes are the commonest adjective complements to the verb *fa*. They include: (*fa*)-*horng* 'reddish', -*hwang* 'yellowish' (orange being divided between the two preceding colors, so far as the *fa*-compounds are concerned), -*liuh* 'greenish', -*ching* 'greenish', -*lan* 'bluish', -*tzzy* 'purplish', *bair* 'whitish, pale', -*hei* 'darkish', -*huei* 'grayish', -*guang* 'glossy' (from the adjective -*guang* 'smooth', to be distinguished from the homophonous and etymologically identical verb-object compound *fa-guang* 'to radiate light'), -*mau* 'rough (to the touch)', -*liang* 'shiny', -*ann* 'darkish', -*lao* 'tough', -*tsuey* 'brittle, (less often) crisp', -*tyan* 'sweetish', -*suan* 'sourish', -*kuu* 'bitterish', -*lah* 'hottish' (rare)—*fa-nenn* 'tender' being also rare, since *nenn* is normally a desirable quality.

2.125. *jaang .de* '(the way a plant or a part of it) grows' can take any adjective as a predicate when the quality or state is taken as the result of growth. Thus, *Jeh shu' jaang .de gau* does not mean 'This tree has grown tall' (= *Jeh shuh jaang-gau.le*) or 'This tree is tall' (= *Jeh shuh gau*), but rather 'This tree is tall (as a result of the way it grows or has grown)'.

2.126. Unlike the complements to *fa*-, which has a slightly pejorative sense, the predicative adjectives after *jaang .de* are of a much wider scope. They include the preceding list as well as others listed below, which are often used in connection with plant growth: (*jaang .de*) *dah* 'large', *sheau* 'small', *charng* 'long' (not necessarily vertically), *doan* 'short', *gau* 'high' (of location), 'tall', *di* 'low' (of location), *ae* 'low, short' (of vertical height), *tsu* 'thick' (in cross section), *shih* 'thin' (cr. sec.), *how* 'thick' (such as leaves), *baur* 'thin' (opposite of *how*), *nenn* 'tender', *lao* 'tough', *jie.shyr* 'strong' (to be distinguished from *jie shyrl* 'to bear fruit' *L*), *feir* 'fat, flourishing', *sheng* 'luxuriant', *wang* 'flourishing'.

2.127. Verbs for man's action on plants have a rather restricted range, if technical words of the botanist, the gardener, and the farmer are excluded. In the nature of the case, they are all transitive verbs. A complication arises when an action is applied to a part of the plant and the whole plant is to be mentioned in the sentence. It is relatively simple when the actor is not mentioned, in which case the plant name is the subject and the action and the part

of the plant takes the verb-object form. Thus, *Tzaodaw hair mei cha-iang .ne* 'The early rice has not transplanted seedlings yet'. If both the actor and the whole plant are to be mentioned, in addition to the action and the part of the plant to which something is done, the usual way is to use the verb *geei* 'give, do something for' before the plant name and make it the first verb in a pair of verbal expressions in series. Thus, *Hualjiang deei geei jey-ke jiu-hual daa-tour .le* 'The gardener will have to clip the (unwanted) buds on this chrysanthemum (plant)'.

2.128. *jonq* 'to plant' (from seed or shoot).

2.128.1. *tzai* 'to plant' (from shoot), 'to transplant'.

2.128.2. *saa-tzeel* 'to scatter seeds'.

2.128.3. *boh-joong(.tzy)* 'to cast seeds' (slightly *L*).

2.129. *jiau* 'to water', as *jiau-hual* 'water the flowers', *jiau-tyan* 'water the field', *jiau-shoei* 'pour on water'.

2.130. *cha-lang* 'to transplant seedlings'.

2.131. *cha-jy* 'to plant a twig'; 'to graft'.

2.132.1. *fen-jy* 'to plant twigs'.

2.132.2. *ia-jy* 'to plant by bending and burying a live branch, so that it will take root'.

2.133. *daa-tour* 'to nip off (unwanted) buds (so that the remaining parts will grow better)'.

2.134. *daa-yaw* 'to spray with germicide'.

2.135. *yah-feir, yah feirlaw* 'to put on fertilizer'.

2.136. *shiou* 'to prune'.

2.137. *tsae* 'to pick' (flowers or leaves).

2.137.1. *chia* 'to nip, clip' (flowers, buds).

2.137.2. *jai* 'to pick' (flowers, fruit).

2.137.3. *jair* 'to pick over' (vegetables).

2.137.4. *shou, daa* 'to reap'.

3. PLANT NAMES

3.0. The following lexicon is arranged according to popular functional ideas of plant classes. The arrangement is roughly that of Ch (see 3.0.1). This is much closer to everyday speech habits than a strictly botanical classification. In order to bring similar words still closer together I have regrouped the headings and entries as follows:

3.1-14. Grains

3.15-23. Beans

3.24-46. Leaf vegetables

3.47-58. Stem and root vegetables

3.59-75. Spicy and odorous vegetables

3.76-88. Fruit and flower vegetables

3.89-93. Mushrooms and fungi

3.94-113. Medicinal plants

3.114-26. Utility plants

3.127-40. Grasses

3.141-85. Flowers

3.186-234. Fruits and nuts

3.235-67. Trees

3.0.1. Identifications of plants are based chiefly on the following works:

Bailey: L. H. Bailey, *Manual of cultivated plants*² (New York, 1949).

BN: *Chih-wu-hsüeh ta-tz'u-tien* [Botanical dictionary] by K'ung Ch'ing-lai et al. (Shanghai, 1920). A number after 'BN' refers to a page in this work.

Ch: *Chih-wu ming-shih t'u-k'ao* [Studies in the names and realia of plants, with illustrations], by Wu Ch'i-chün (editor's preface T'aiyuan, 1848; reprinted Shanghai, 1936).

Lf: Berthold Laufer, *Sino-Iranica* (Field Museum publications, Anthropol-

logical series 15.185-630, 1915-17). A number after 'Lf' refers to a page in this work.

Rd: Bernard E. Read, *Chinese medicinal plants from the Pen ts'ao kang mu A.D. 1596 [originally by Li Shih-chên]*⁵ (Shanghai, 1936). A number after 'Rd' is the serial number, not a page number.

CHPT: Bernard E. Read, *Famine foods listed in the Chiu huang pen ts'ao* (Shanghai, 1946).⁶ Numbers before and after a period are respectively section and plant numbers.

3.0.2. Ch gives only Chinese names, but its copious excellent illustrations have been a great help to the writer in identifying Chinese names. Rd gives popular English names of plants as well as Latin names. These English forms have been followed in the main, but if an English name is found only in Rd and not otherwise known to the writer, it is prefixed by 'Rd' (e.g. in No. 99). A literal translation of a compound is prefixed by 'lit.' (e.g. No. 83), but if it also has a regular English equivalent, then the literal form will only be followed by 'i.e.' and the regular English (e.g. No. 46).

3.0.3. For botanical identification, if the authorities cited agree, then the Latin name will be given once after the reference numbers, otherwise separately. (Rd usually notes variant identifications by other authorities.) I have included a few very recent importations, which nevertheless are frequently mentioned in China now.

(3.)1.⁷ *may.tzy*, *sheau₀may* 'wheat'. Ch 11. Rd 764 *Triticum vulgare*, Vill., BN 87 *T. sativum*, Lam. var. *vulgare*, Hack. *may-suell* 'ear'; *miann F*, *miannfeen* 'flour'; *fu.tzy*, *fu.pyi*, *mayfu(.tzy)* 'bran'; *miann.jin* 'flour-muscle, i.e. gluten'; *mayya(l)* 'malt'.

2. *daw F* 'rice' (the plant or unhulled grain). Ch 16. Rd 745, BN 1228 *Oryza sativa*, L. Lf 372 gives no Latin name. *mi F* '(hulled) rice'; *fann F* '(cooked) rice'; *jou F* 'gongee, i.e. thin boiled rice'; *miifeen* 'rice flour'.

2.1. *shianmli* 'common rice'. Rd 747, BN 1229 O.s., L. var. 'Annamese upland rice.'

2.2. *nuoh₀mli* (dial.) 'glutinous rice', *jiang₀mli* 'river rice, i.e. glutinous rice'. BN 1521 O.s., L., var. *glutinosa*, Mats. *nuoh₀mi-fann* (dial.), *jiang₀mli-fann* 'glutinous rice (cooked)'.

3. *dahmay* 'large wheat, i.e. barley'. Ch 12. Rd 741 *Hordeum vulgare*, L., BN 54 *Hordeum sativum*, Jess. var. *vulgare*, L. *mayya(l)* 'malt'.

4. *sheau₀mli*, *sheau₀mleel*, (*suh L*) 'small rice, i.e. (short) millet'. Ch 10. Rd 760, BN 1090 *Setaria italica*, Kth. var. *germanica*, Trin.

5. *sheau₀mli*, *sheau₀mleel*, (*liang L*) '(spiked) millet, (Italian) millet'. Ch 13. Rd 758. *Setaria italica*, Beauv. vel Kth.

6. *shuu.tzy* (*shuu L*) 'glutinous millet'. Ch 14. Rd 752 *Panicum miliaceum*, L. var. *glutinosa*, Bretsch.

(7. *jih L*. Ch 15. Rd 751 'panicked millet, common millet, Indian millet', BN 1346 *Panicum miliaceum*, L. See also discussion of term in Ch 25-30.)

8. *gau.liang* (*shuushuu*, *shuushwu L*) 'sorghum, kaoliang'. Ch 24. Rd 730, BN 1231 *Andropogon sorghum*, Brot. var. *vulgaris*, Hack. The term is also applied to the colorless spirit distilled from the grain. The stalk is commonly called *shwu.ji* or *shwu.jie*, but Rd 731, BN 1516 regard *shwu.jie* as var. *saccharatus*, Koerm., as it is often eaten like sugar cane. Also *gau.liang-gaan.tzy* 'sorghum stalk'.

⁵ The original work, c. 1400, was by Chu Su. Read apparently took the author's enfeoffment title 'Chou Ting-wang' for his name, Chu Su being the fifth son of the first emperor of the Mings.

⁷ For simplicity the main section number '3' will not be repeated after this.

9. *lao-yuh.mil* 'old jade-rice', *banq.tzy* 'club-suffix' (from shape of the ear)—i.e. '(Indian) corn, maize'. Ch 40. Rd 765, BN 278 Zea mays, L. Name also applied to ear or grain. *banq.tzy-miall* (compound of two suffixed words) 'corn meal'. The Nanking form *yihchü.chu* is popularly regarded there as a compound with some meaningless first element *yih-* and a reduplicated expressive word *chü.chu*. Actually it is nothing but the usual *L* name *yuh-shuushuu* 'jade Szechwan millet' for Indian corn. Since one type of Nanking pronunciation has *i* for *ü*, with optional falling tone on the word for 'jade', optional affrication in the word for 'Szechwan' (*ch* for *sh*), and required affrication in the word for 'millet', the resulting form *yihchü.chu* is phonologically perfectly regular.

(10. *yannmay*, *yannmoh* *L* 'oats'. Ch 38. Rd 733 a. Avena fatua, L., b. A. sativa, L., BN 792 A. f., L. The writer was not personally acquainted with this grain in China, except in the form of recently imported American rolled oats under the name of *maypyi*, *maypyel* 'wheat skin', or, in cooked form, *mayjou* 'wheat-congee, i.e. oatmeal'. Ch 38 quotes CHPT as saying that oats can be used for making congee, but of inferior quality.)

11. *chyau.may* 'buckwheat'. Ch 22. Rd 564, BN 1383 Fagopyrum esculentum, Moench. *chyau.may-miann* 'buckwheat noodles'.

12. *bay.tzy* Rd 'barnyard grass'. Ch 32. Bailey 148, Rd 749, BN 1226 Panicum crus-galli, L.

13. *jy.ma* (*hwuma* *L*) 'sesame', Rd also 'teel'. Ch 6. Rd 97, BN 674, Lf 289 Sesame indicum, L. *mayou*, *shiang-you* 'sesame oil'; *jy.majiang* 'sesame jam', also, more rarely, *majiang*.

14. *yihren(mil)*, (*yihyii* *L*) Rd 'Job's tears'. Ch 8. Rd 737, BN 1435 Coix lachryma, L.

15. *dah-dow* 'large bean', *maudow*(l) 'hairy bean', *hwangdow*(l) 'yellow bean'—i.e. soybean'. Ch 9. Rd Glycine soja, S. et Z., 388 pekinensis, 389 flava, 390 alba, BN 47 Glycine hispida, Moench., 1133 G.h., Maxim. *dowyal(tsay)* 'bean sprouts'; *dow(fu)jiang* 'bean milk'; *dow.fu* 'bean curd'; *dowyou* 'bean oil'; *jiang F* 'soybean jam', Rd 389c 'soybean paste'; *jiangyou* 'jam oil, i.e. soybean sauce'; *dowchyy*, *dowshyr* 'salted fermented soybeans', Rd 388b 'bean relish'. It should be noted that in all the compounded names for beans and peas with *-dow*(l) the unsuffixed forms are more common than the suffixed forms, but that for the GENERAL word, only the suffixed form *dowl* is possible.

16. *tsarndow*(l) 'silkworm bean, i.e. horse bean, fava bean'. Ch 23. Rd 413, BN 1577, Lf 307 Vicia faba, L.

17. *beandow* lit. 'flat bean', Rd 'hyacinth bean, Egyptian kidney bean'. Ch 13. Rd 382, BN 1461 Dolichos lablab, L.

18. *yunbeandow*. Ch 43. Unid. Term also used for the pods of *beandow* used as a green vegetable.

19. *uandow*(l) 'peas'. Ch 40b. Rd 402, BN 1359, Lf 305 Pisum sativum, L. *uan.dow-miaul* 'peashoots' (used as vegetable).

20. *liuhdow*(l) 'green bean, i.e. mung bean'. Ch 21. Rd 400, BN 1287 Phaseolus mungo, L. var. radiatus, Bak. *dowyal*, *liuh.dowyal*, *chiatsay* 'clipped vegetable, i.e. peasprouts, (mung)bean sprouts'.

21. *jiangdow*(l) 'cowpeas, black-eyed bean'. Ch 40a. Rd 417, BN 836 Vigna sinensis, Hassk.

22. *daudow* 'knife bean, i.e. broad bean'. Ch 41. BN 19 Canavallia ensiformis, DC.

23. *horng-dow*(l) lit. 'red bean', (*chyh-sheaudow* *L* 'red-small-bean'). Ch 8a. BN 475 Phaseolus mungo, L. var. Subtriloba, Fr. et Sav.

24. *bair-tsai* 'white vegetable', *hwangya*(l)-*tsai* 'yellow-sprout vegetable' (second name sometimes limited to the kind bleached yellow from being kept away from light), (*song* *L*)—i.e. 'celery cabbage'. Ch 67. Rd 476 Brassica pekinensis, Rupr., BN 1059 Brassica chinensis, L. var.

25. *youtsai* 'oil vegetable', *ching-tsai* 'green vegetable', *bair-tsai* 'white vegetable'—i.e. 'Chinese green, bok choy' (the latter being the Cantonese cognate for *bair-tsai*, used since 1950 at a Berkeley, Calif., chain market), Rd: 'Chinese colza(?)'. Ch 83. BN 1388 Brassica campestris, L., Rd 473 var. oleifera, DC., Lf 381 B. rapa. *tsayyou* 'vegetable oil (from this vegetable)'; *tsaytair*, *tzyy-tsaytair* 'vegetable top' (usually purple kind preferred), term applied to the budding or near blossoming top part, or to the plant specially cultivated for this part, for which the variety in Hupeh is best known.

- 25.1. *sheau-bairtsay*, -bairtsall a small variety of No. 25, common in southern Kiangsu.
26. *yang-bairtsay* 'foreign cabbage', *bautsay* 'wrapped cabbage'—i.e. 'cabbage'. Bailey 436 *Brassica oleracea capitata*, L.
27. *tsayhual* 'cauliflower' AN -geh (not -duoo!), -ke. Bailey 436 *Brassica oleracea botrytis*, L.
28. *jiehtsay*. Ch 68(?). Unid.
29. *chuen.bulao* '(in) spring not tough', *sheue.lil-horng* 'in snow (it's) red',—i.e. 'mustard green.' Ch 68(?). Rd 474(?) Lf 380 *Brassica juncea*, Coss., BN 556 *B. cernua*, Thunb. (Under *jieh* Rd has *ching-jieh* and *tsyhjieh* 'black mustard'.)
30. *bo(r)tsay*, (*boleng* L) 'spinach'. Ch 93. Rd 563, Lf 393 *Spinacia oleracea*, L., BN 1069 S. o., Mill.
31. *shianntsay* Rd 'amaranth', CHPT 'wild amaranth'. Ch 51. Rd 557 (a. *bair-shiann*), BN 939 *Amarantus mangostamus*, L., CHPT 13.25 *A. blitum*, L.
- (31.1. Rd 557b. *horng-shiann*, BN 1127 *yannlai-horng*, lit. '(when) wild geese come (then) red', a garden flower, *Amaranthus gangeticus*, L. The writer has eaten of reddish *shianntsay*, which apparently is not No. 31.1, as described in BN, but still No. 31.)
32. *maa.shy-shiann* (<*maachyy-shiann*, lit. 'horse-tooth amaranth') Rd 'purslane'. Ch 52. Rd 554, BN 843 *Portulaca oleracea*, L.
33. *wenqtsay*, *kongshin-tsai*, lit. 'hollow-center vegetable'. Ch 95., Rd 157, BN 1439 *Ipomoea aquatica*.
34. *jihtsay* 'endive', Rd 'shepherd's purse'. Ch 65. Rd 478 *Caspella burs-pastoris*, L. BN 1458 & CHPT 14.27 Mœnch. [zia²ts'e²] in Wu.
35. *maalantour* CHPT 2.3 *Aster trinervius*, L.
36. *chyntsay*, (*shoeichyn* L) '(Chinese tubular) celery'. Ch 63. Rd 225, BN 228 *Oenanthe stolonifera*, DC.
37. *yawchyn* 'medicinal(-odored) celery', *yang-chyntsay* 'foreign celery'—i.e. 'celery'. Bailey 755, BN 511 *Apium graveolens*, L.
38. *tornghau(tsay)*, *pernghau*, *hau.tzy* Rd: '(garden) daisy', but CHPT 13.31 'garland chrysanthemum'. Ch 92a. Rd 23, BN 802 *Chrysanthemum coronarium*, L. The Chinese variety, probably different from the usual garden variety, is cultivated for its tender edible leaves.
39. *pyau.erl-tsai*. Unid.
40. *täkuutsay*. Unid. Numbers 39 & 40 seem to be varieties of celery cabbage common in southern Kiangsu, with very flat head and dark green leaves. The first syllable in No. 40 is in the entering tone, short, but without glottal stop in non-final position. The writer would be at a loss as to what tone to give it in Mandarin.
41. *chwentsay*, a jelly-covered water vegetable grown in Hangchow. Ch 425. Rd 540 'water shield' or 'water mallow' *Brasenia peltata*, Bursh., BN 1333 *Brasenia purpure*, Casp.
42. *lusoan* 'reed-shoot', *longshiu-tsai* 'dragon-beard vegetable',—i.e. 'asparagus'. BN 823 *Asparagus officinalis*, L.
43. *shiyang-tsai* 'occidental vegetable, i.e. water-cress'. Bailey 447 *Nasturtium officinale*, R. Br., *Rorippa nasturtium-aquaticum*, Hayek.
44. *tzyy-tsai* 'purple laver'. Ch 426. Rd 865 *Porphyra laciniata*, Ag., BN 1106 *P. tenera*, Kjellm.
45. *haeday* 'sea ribbon, i.e. kelp'. Ch 427. BN 760 *Laminaria*, Rd 863 *L. religiosa*, Miyabe.
46. *sheng-tsai* 'raw vegetable, i.e. lettuce', term applied to imported lettuce cultivated for its leaves. Rd 39a *Lactuca scariola*, L. var. *sativa*, Bisch.
47. *uo.jiu*, *uosoan* 'lettuce'. Ch 89. Rd 39, Lf 401 *Lactuca sativa*, L., BN 1207 L.s., Bisch. Cultivated in China for the fleshy stem (hence -soan) rather than the leaves.
48. *jiaubair*, *jiaueel-tsai*, *jiaultsay*, (BN *gu*) Rd: 'water bamboo, Indian rice'. Rd 766, BN 1065 *Zizania aquatica*, L.
49. *plee.la* 'kohlrabi' AN -geh, -ke. Rd 475, BN 287 (*ganlan*) *Brassica oleracea*, L.
50. *byi.chi* (*fwutsyh*, *u-yuh* 'black taro' L) 'water chestnut' (to be distinguished from No. 211), AN -geh. Rd 727 *Scirpus tuberosus*, Roxb., BN 790 *Heleocharis plantaginea*, R. Br.

51. *luo.bo, luo.bei* (*laiboh* L) 'radish' AN -*gen*, -*gel*, -*ke*. Ch 84b. Rd 482, BN 1061 *Raphanus sativus*, L. *luo.bo-ing.tzy* 'radish top'.

51.1. *horng-luo.bo* 'red radish'.

51.2. *bair-luo.bo* 'white radish'.

51.3. *say-sheuelli* 'excels snowpear', *luo.bo-say-sheuelli* 'radish excels snowpear', a stocky large greenish white radish, cultivated in the north as a winter fruit. The name probably came from a street cry, of which it is a part.

52. *hwuluo.bo, hwuluo.bei* 'foreign radish, i.e. carrot' AN -*gel*, -*ke*. Rd 219, BN 680, Lf 451 *Daucus carota*, L.

53. *bae.her* 'lily'. Ch 58. CHPT 6.2, Rd 682, *Lilium brownii*, Spae, 682a *Lilium tigrinum*, Gawl. 'Peking species (tiger lily)', BN 388 L. *japonicum*, Thunb. Name applied to the pl or edible bulb, but *bae.her-hual* for the flower. AN for root -*geh*, for slice -*ball*.

54. *bair-shuu* 'white potato', *horng-shuu* 'red potato', (*shan-yuh* 'mountain taro' dial.), (*ganshuu* 'sweet potato' L, Rd *bair-shuu* 'white potato')—all are variant names (irrespective of actual color) for 'sweet potato' AN for root -*geh*. Ch 124. Rd 158, BN 287 *Ipomoea batatas*, Lam.

55. *shan.yaw* 'mountain medicine, i.e. (cylindrical) yam' AN for tuber -*gen*, -*gel*. Ch 57a (*shuuyuh* L). Rd 657, BN 1456 *Dioscorea japonica*, Thunb., Rd 657, *Dioscorea batatas*, Dene. To be distinguished from 'yam' which in some parts of U. S. means 'sweet potato' or 'dark brown sweet potato'.

56. *shan.yaw-dall* 'mountain-medicine egg', *tuudowl* 'earth-bean', (*yang-shanyuh* 'foreign sweet-potato' dial., *maaling-shuu* 'horse-bell potato' L)—i.e. '(Irish) potato' AN for tuber -*geh*. BN 846 *Solanum tuberosum*, L.

57. *yuh.tour* 'taro' AN for tuber -*geh*. Ch 80. Rd 710, BN 612 *Colocasia antiquorum*, Schott.

58. *tsyr.gu* (Rd *her-utsy* 'river waterchestnut', the top called *jeandau-tsao* 'scissors-grass' *yannwoei-tsao* 'swallowtail grass') 'arrowhead' AN for tuber -*geh*. Ch 703a. BN 1257 *Sagittaria sagittifolia*, L., Rd 781 var. *sinensis*, Mak.

59. *lah-jiau, chinjiau, lah-chinjiau* 'hot (green) pepper, hot (red) pepper'. Ch 131. BN 1077 *Capricum longum*, L. The form *chinjiau* is probably a blend of *chynjiau* (cf. No. 61) and *ching-jiau* 'green pepper'. In the central dialects *ching* and *chin* are both pronounced *chin*.

60. *yang-lahjiau, yang-chinjiau* 'sweet pepper'. BN 1174 (*shytour fanjiau* 'lionhead foreign pepper') *Capsicum annuum*, L. var. *grossum*, Seudt.

61. *huajiau* lit. 'flower pepper', *chuanjiau* lit. 'Szechwan pepper' (*chynjiau* 'Chyn (i.e. NW region) pepper' L, *shuujiau* 'Szechwan pepper' L). Ch 727. Rd 360 'Japanese pepper', BN 827 *Xanthoxylum* (*Zanthoxylum*) *piperitum*, DC., Rd 358 'Chinese pepper or fagara', BN 396 *Xanthoxylum* (Z.) *alatum*, Roxb.

62. *hwujiau* 'foreign pepper, i.e. (black pepper'. Rd 631, BN 676, Lf 375 *Piper nigrum*, L. *hwujiau*-(*miall*) '(ground) pepper'.

63. *bajeau, bajeau-hweishiang* 'eight-cornered fennel, i.e. star anise, Chinese anise'. Rd 506 BN 18 *Illicium verum*, Hook.

64. *jiang F, sheng-jiang* 'raw-ginger, i.e. ginger' AN for root -*kuay*. Ch 59. Rd 650, BN 1435, Lf 545 *Zingibar officinale*, Rose.

65. *tsong F* 'spring onion, scallion'. Ch 60. Rd 666, BN 1211 *Allium fistulosum*, L.

66. *yang-tsong* (*hwutsong* L) 'foreign onion, i.e. onion'. Rd 664, BN 277 *Allium cepa*, L.

67. *suann F, dah-suann* (but Ch *sheau-suann*) 'garlic' AN for section -*yal*. Ch 76. Rd 671, Lf 302 *Allium sativum*, L. *suannmiall* 'garlic shoots';

(68. *hwu* L (?). According to Ch 76 *Sheau-suann wei suann, dah-suann wei hwu* 'small garlic is garlic, large garlic is *hwu*'. Rd equates *hwu* with 'chive'. Rd 672, BN 1209, Lf 302 *Allium Scorodoprasum*, L.)

69. *jeou.tsay* 'leek' (very thin-bladed variety). Ch 73. Rd 670, BN 721 *Allium odorum*, L. *jeouhuwang*(l) 'tender leaves bleached yellow from darkness'.

70. *hwei-shiang* (*hwaishiang* L) 'fennel'. Ch 84. Rd 222 *Foeniculum vulgare*, Mill., BN 1568 F. *officinale*, All.

71. **shiang-tsay** lit. 'fragrant vegetable', (*y(u)an-suei-tsay*, dial.), (*hwusuei* L), i.e. 'coriander'. Ch 92. Rd 217, BN 673, Lf 297 *Coriandrum sativum*, L. In America often called 'Chinese parsley'.

72. **ding.shiang** (*dingtzyy-shiang* L) 'cloves'. Rd 244 *Eugenia caryophyllata*, Willd., BN 5 *Jamlosa caryophyllus*, Ndz. \neq 161.

73. **charshuh** (*ming* L, *kuu-twu* L) 'tea (tree)'. Ch 754. Rd 267, CHPT 9.1, BN 800 *Thea sinensis*, L. *charyeh* 'tea (leaf, as household article)'; *char* a. 'tea (the infused drink)'; b. 'tea (leaf, as trade commodity)'; *ching-char* 'green tea' (leaf or drink); *horng-char* 'red tea, i.e. black tea' (leaf or drink).

74. **kafel, jlafel** 'coffee'. BN 512 *Coffea arabica*, L.

75. **shiang-tsao** 'fragrant grass, i.e. vanilla grass'. Rd 740 (*bairmau-shiang*), BN 696 (*maushiang*) *Hierochloa borealis*, Roem et Sch.

76. **dong.gua** 'winter melon, vegetable marrow'. Ch 57. Rd 56 ('white gourd, gourd melon'), Bailey 953 *Benincasa hispida*, Cogn. (BN 254 *Benincasa cerifera*, Savi.)

77. **huh.tzy** '(summer) squash' (?) Ch 84a. Unid.

78. **hwu.lu** 'bottlegourd'. Ch 63. Rd 62, BN 1272, Lf 197 *Lagenaria vulgaris*, Ser.

79. **uo.gua, nan.gua** (Rd 61 also *shi-hwu.lu*, which in the writer's speech is applied to a more elongated melon), ([pökwo] 'north-melon', [vekwo] 'foreign melon' Wu dial.) 'pumpkin'. Ch 126a. Rd 61 BN 626 *Cucurbita pepo*, L. *uo.gua-tzeel* 'pumpkin seed'.

80. **sy.gua** lit. 'fiber melon' Rd 'vegetable sponge or loofah'. Ch 126b. Rd 64, BN 1092 *Luffa cylindrica*, Roem. The tender melon is used as vegetable. *sy.gua-law(.tzy)* (dial.) 'melon sponge' (of ripe melon, for scouring dishes). The Standard form, which the writer does not use, is *sy.gua-rang.tzy*.

81. **tsay.gua** lit. 'vegetable melon'. Ch 90 (*yuehqua* L). Rd 59 *Cucumis melo*, L. var. *conomon*, Mak., BN 1123 *C. conomon*, Thunb.

82. **hwang.gua** 'yellow melon' (*hwugua* 'foreign melon' L)—i.e. 'cucumber' AN for melon -*gel*. Ch 96. Rd 60, BN 670, Lf 300 *Cucumis sativus*, L.

83. **iyegua** lit. 'festival melon'. Unid. A Cantonese vegetable melon planted between the fifth and the eighth moon (i.e. between two major festivals). Also called *maugua* 'hairy melon'.

84. **kuu-gua** 'bitter melon', acc. to Rd = 'wild cucumber'. Ch 101. Rd 65, BN 685 *Momordica charantia*, L.

85. **chye.tzy**, (*luohsu*, dial.) 'eggplant'. Rd 119, BN 695 *Solanum melongena*, L.

86. **jitour-mil** lit. 'chicken-head rice,' (*chiannshyr* L). Ch 697. Rd 541 'chicken head, foxnut', BN 559 *Euryale ferox*, Salisb.

87. **jin.jen(-tsay)** lit; 'gold-needle (vegetable)', **hwanghual-tsai** lit. 'yellow-flower vegetable', (*shiuntsao* L), Rd 'yellow day lily'. Ch 342. Rd 679, BN 36 *Hemerocallis fulva*, L. The dried flower is used as vegetable, often together with No. 92.

88. **shl.horng-shyh** 'occidental red persimmon', (*fanchye* 'foreign eggplant' dial.)—i.e. 'tomato'. BN 713 *Lycopersicum esculentum*, Mill.

89. **mo.gu, shian-mo, shin.shian mo.gu** 'fresh stubby mushroom' (Rd *moqu-jiunn*, *rowjiunn*). Rd 829 *Clavaria pistillarius*, L.

89.1. **koou-mo** '(Jang.jia-)koou mushroom, i.e. dried Kalgan mushroom', also 'any dried stubby mushroom'.

90. **shiang-jiunn, shian-jiunn** '(thin top) mushroom'. Rd 824 *Agaricus breitschneideri*, Kalich. et Thün, BN 732 *Cortinellus shiitake*, Henn.

90.1. **shiang-jiunn, shiang.gu, donggu**, terms applied to dried form of No. 90.

90.2. **tsaogu, huagu**, varieties of above, commonly used in dried form.

91. **jitong(-jiunn)**. Rd 828 *Clavaria corniculata*, Schaeff. This is a large meaty mushroom, much cultivated in Yunnan.

92. **muh'eel**, lit. 'wood-ear, tree-ear'. Rd = 'Job's ear'. Rd 827A, BN 186 *Auricularia auricula-judae*, Schr.

93. **shyr'eel**, lit. 'rock-ear'. Rd 818, BN 325 *Gyrophora rellea*, Ach.

94. *renshen*, *shen* *F* 'ginseng'. Ch 140. Rd 237, BN 13 *Panax ginseng*, C. A. Mey. *shenshiu* 'root hair of ginseng'.

95. *hwangchyi* Rd 'yellow vetch'. Ch 141. Rd 372 *Astragalus hoantchy*, Franch., BN 193 *A reflexistipulus*, Miq.

96. *gantsao* lit. 'sweet grass' Rd 'licorice'. Ch 143. Rd 391, BN 284 *Glycyrrhiza glabra*, L.

97. *bair-jwu*, (BN also *tsang-jwu* L). Ch 144. Rd 14, BN 1273 *Atraclylis ovata*, Thunb.

98. *farnfeng* lit. 'ward-wind'. Ch 150. Rd 233, BN 450 *Siler divaricatum*, Bth et Hk.

99. *hwang-lian* Rd 'golden thread'. Rd 534, Lf 546 *Coptis teeta*, Wall., BN 1141 *Coptis japonica*, Makino. Word often used as an example of extreme bitterness, lit. or fig.

(100. *jiann-joongshiau* '(plant) meets swelling (and swelling) subsides'. The writer does not know any herb under this name and cites it for the interesting grammatical construction. Rd 256 identifies it with *Viola mirabilis*, L., but Ch gives two different plants 201 and 348 under this name.)

101. *dong-chorng-shiah-tsao* (-*jiunn*) 'winters-(it's)-insect, summers-(it's)-grass (-plant)'. Ch 230. BN 257 *Cordyceps robertii*. A Yunnan plant which grows as parasite in a caterpillar host, popularly believed to be one organism shifting between insect and plant. Used, in dried form, as a medicinal vegetable.

102. *ma₂hwang* 'ephedra'. Ch 261. Rd 783 *Ephedra sinica*, BN 1004 *Ephedra vulgaris*, Rich. var. *helvetia*, H. et T.

103. *ay F*, *chyl* (jou)-*ay* Rd 'moxa or mugwort'. Ch 273. Rd 9, BN 386 *Artemesia vulgaris*, L. var. *indica*, Maxim.

104. *day.hwang* 'great yellow, i.e. rhubarb'. (Note special pron. of word for 'great'.) Ch 553. Rd 582, BN 57 *Rheum officinale*, Baill., Lf 549 *R. ribes*. Used in China as laxative only.

105. *bihma* 'castor oil plant'. Rd 331, BN 1275, Lf 403 *Ricinus communis*, L. *bihma* (tzy) -*you* 'castor oil'.

106. *bann-shiah*, lit. 'half-summer'. Ch 564. Rd 711, BN 261 *Pinellia tuberifera*, Ten.

107: *dangguel*, lit. 'ought to return'. Ch 583. Rd 210 *Angelica polymorpha*, Maxim. var. *Sinensis*, BN 1221 *Ligusticum acutilobum*, S. et Z.

108. *dowkow* 'nutmeg'. Ch 588. Rd 642 (*dowkow*, *tsao-dowkow*) *Amomum* 'Chinese cardamon', Rd 643 (*tsao-dowkow*) *Amomum globosum*, Lour. 'wild cardamon', BN 813 *Amomum costatum*, Roxb. (*tsao-dowkow*, namely *dowkow*). ≠ No. 252.

109. *sharen*, *sharel*, (*suoshamih* L) Rd 'bastard cardamon'. Rd 644, BN 1444, Lf 481 *Amomum xanthioides*, Wall.

110. *boh.he* 'mint', Rd 'wild mint'. Ch 593. Rd 129 *Mentha arvensis*, L., BN 1432 var. *piperascens*, Holmes. The Wu dial form [buhu] instead of [bōwu] agrees with an ancient reading *b'uā²-xā²*.

111. *huooshiang* (a kind of mint with pointed leaves). Ch 599. Rd 128 *Lophanthus rugosus*, Rd Fisch., BN 154 L. r., Fisch. et Mey.

112. *ing.suh.hual*, (*ingtzyy-suh* L) 'poppy'. Ch 620. Rd 490 *Papaver somniferum*, L., BN 1522 var. *nigrum*, DC. *iapiana* (~*yea*-), *ia.piann-ian* (~*yea*-), *dah-ian* 'opium'; *iantuu*, *tuu F* 'opium' (prepared in lump form).

113. *yun₂shiang* 'rue'. BN 573 *Ruta graveolens*, L. Name also applied to the resin-like essence of rue.

114. *dah-ma*, *ma.tzy* 'hemp'. Ch 7. Rd 598, BN 55 *Cannabis sativa*, L. *ma.tzy* 'hemp seed'; *ma F* 'hemp' (the fiber); *mashengl* 'hemp twine'; *masheng* 'hemp rope'.

115. *juhma* Rd 'ramie, grass-cloth plant'. Ch 328. Rd 592, BN 694 *Boehmeria nivea*, Hk et Arn.

116. *gee*. Rd 'ke hemp'. Ch 507. Rd 406 *Pueraria hirsuta*, Schneid., BN 1213 *Pueraria thunbergiana*, Bth.

117. *mian.hua* 'cotton plant'. BN 815 (*tsaomian*) *Gossypium herbaceum*, L. Name also applied to 'cotton'; *mian.hua-tzeel* 'cotton seed'; *mian.hua-you* 'cotton oil'.

118. *bairjl*, *bairjyl*. Ch 168. Rd 634 *Bletia hyacinthina*, R. Br., BN 295 *Bletilla hyacinthina*, Rehb. F. Term also applied to the root, used for grinding vermilion.

119. **ternghwang**. Rd 'gamboge' 265, BN 1491 *Garcinia morella*, Desv. Term also applied to the yellow pigment from the juice.

120. **tongtsao**, (*tongtuomuh* L) 'rice paper plant'. Ch 186. Rd 238 *Fatsia papyrifera*, Bth et Hk, BN 943 *Arabia papyrifera*, Hook. Name also applied to the pith (rice paper), used for making delicate ornaments.

(121. **tongtsao**, *wannnian-tern* lit. '10,000-year vine'. Rd 522, BN 942 *Akebia quinata*, Dene. ≠ No. 120. Plant not known to writer.)

122. **dengshin-tsao** 'lamp-wick grass, i.e. common rush, soft rush'. Ch 340. Rd 696 *Juncus effusus*, L., BN 1381 var. *decipiens*, F. *dengtsao*, (*dengshin* dial.) 'lamp-wick' (made from the pith).

123. **woei.tzy**, *luwoei* 'common reed'. Ch 331. Rd 754 *Phragmites communis*, Trin., BN 1515 var. *longivalvis*, Miq. *luochair* 'rush fuel' (made of the stalks); *shyi.tzy*, *shyi* F, *lushyi* 'matting' (made from the splints); *dyi.moll* 'flute membrane' (made of the inner skin, which is stretched over a special hole of the Chinese flute to give it its characteristically reedy quality).

(124. *ruoh* L Ch 352. BN 102 *shanbairjwu* *Bambusa vertchii*, Carr. The form *ruoh* corresponds to the Wu dial. form *ná?*, which occurs in the word *tzongná?*, for the leaves of this plant used for wrapping tetrahedrons of glutinous rice, called both in Mand. and Wu *tzong.tzy*, alternating, in Mand. only, with *jonq.tzy*.)

125. **terng.tzy** 'rattan'. BN 701 *Calamus rotang*, L. Word also used for vines in general.

126. **jwu.tzy** 'bamboo'. Rd 755, BN 13, 377, 1348, 1419, etc. *Phyllostachys* (several species, esp. *P. mitis*, Riv.). *soen* F 'bamboo shoot'. Two of the commonest kinds are *dongsoen* 'winter shoot' (of the short, thick kind); *ching-soen* 'green shoot' (thin, long, usually salted kind).

127. **chiu.chuel-tsao** lit. 'cricket grass', so-called because it is used for stirring crickets to a fight, probably = 'crab grass'. *Digitaria sanguinalis*. Rd 143 equates D. s. to *Caryopteris divaricata*, Max., which, from the illustration in BN 1389, is not *chiu.chuel-tsao*.

128. **goou-yil.ba-tsao** lit. 'dog-tail grass', (*yow* L) Rd 'green foxtail'. Ch 281. Rd 761a, BN 530 *Setaria viridis*, Beauv.

129. **pwugong-ing**, *par-eel.dou-tsao* 'scoop-ear-grass' (from resemblance of the thistle to an ear-scoop)—i.e. 'dandelion'. Ch 334. Rd 48 *Taraxacum officinale*, Web., BN 1270 T. o., Wigg., var. *glaucescens*, Koch.

130. **changpwu** 'calamus', Rd also 'sweet flag'. Ch 420. Rd 703 *Acornus calamus*, L., Rd 704 *Acornus gramineus*, Ait., BN 301 (*bair-chang*, *shyr-changpwu*) *Acorus calamus*, L.

131. **parshan-huu**(1) 'climb-mountain tiger', **parchyang-huu**(1) 'climb-wall tiger' (*charng-chuen-tern* 'ever-spring vine' L)—i.e. '(common) ivy'. Ch 455. Rd 239 *Hedera helix*, L., BN 886 var. *colchica*, C. Koch.

132. (Rd *gouwoen* 'hooked lips', *dwu-gen* 'poisonous root', *hwang-tern* 'yellow vine'; BN *yee-gee*,—all L—i.e. 'poison ivy'. Rd 317, BN 984 *Rhus toxicodendron*, L. var. *radicans*, Miq. The writer has never met with this plant in China nor heard it spoken of in Chinese.)

133. **fengwoei-tsao** 'phoenix-tail grass, i.e. fern'. Ch 386, 387. Rd 800 ff, order of Filicales. Rd 811, BN 1385 (*jyue* L) *Pteridium aquilidum*, Kuhn.

134. **harnshiou-tsao** 'with-bashfulness grass, i.e. sensitive plant'. BN 436 *Mimosa pudica*, L.

135. **woasong** 'tile pine, i.e. roof pine'. Rd 469 *Cotyledon fimbriata*, Turcz var. *ramosissima*, Maxim., BN 645 *C. japonica*, Maxim.

136. **shianren-jaang** 'immortal-man palm (of the hand), i.e. cactus'. Ch 366. BN 251 *Opuntia* (subgenus *Platyopuntia*). The combining form *shianren-* is also applied to other cacti characterized and named by *-chyou* 'ball', *-bang* 'pole' (*O. cylindrica*, DC), etc.

137. **wannnian-ching** '10,000-year green, i.e. Chinese evergreen'. Ch 367. BN 1194 *Rhodea japonica*, Roth.

138. **pyngtsao**, *fwupyng*, *shoelpyng* Rd 'duckweed'. Ch 422. Rd 702 *Lemma minor*, etc., BN 221 *Spirodela polyrrhiza*, Schleid. Ch 422a *pyntsao* is often confused with the above, especially as *pyn* and *pyng* are pronounced alike in the central dialects.

- 138.1. *shoeltsao* (*tsao* *L*) lit. 'water-grass', the kind used for keeping goldfish. Unid.
139. *ching*(*ni*)-*tal*(*r*) 'green(mud) moss, i.e. moss'. Class Musci, e.g. BN 1248 *Mnium punctatum*, Hedv.
- (140. *Wang-buh-liou-shyng* lit. 'The king does not keep (you) from going' Rd 'cowherb'. Ch 271. Rd 551 *Saponaria vaccaria*, L. BN 246 *Vaccaria vulgaris*, Host.)
141. *lanhual* '(Chinese) orchid'. Ch 612. BN 1535 *Cymbidium virens*, Lindl. *jiannlan* '(Fu)kien *lan*'; *suhshin-lan* lit. 'plain-centered *lan*'; *huenshin-lan* lit. 'rich-centered *lan*', the former with greenish petals and greenish pistil and stamina, the latter with warm colors on the pistil and stamina.
- (142. Ch 640 *lanhua-shuangyeh-tsao* '*lan*-flower-doublepetal plant, i.e. (occidental) orchid'. BN 1535 *Cypripedium japonicum*, Thunb. The writer knows of no popular name except descriptive phrases like *way.gwo lan-hual* 'foreign *lan*-flower'.)
143. *shaur.yaw*(*hual*) 'Chinese peony'. Ch 584. Rd 536, BN 464 *Paeonia albiflora*, Pall.
144. *muu.dan*(*hual*) 'tree peony'. Ch 584a. Rd 537 *Paeonia moutan*, Sims, BN 461 P. m., Ait.
145. *herhual*, (*lianhua*, *lian'ouo* *L*) 'lotus'. Rd 542, BN 1331 *Nelumbo nucifera*, Gaertn. *ouo* *F* 'rootstock of lotus' (eaten as fruit or vegetable); *heryeh* 'lotus leaf'; *lian.perng* 'lotus fruit'; *liantzzy*, (*lianshin* dial.) 'lotus seed'; *ouuseen* 'lotus starch' (made from *ouo*, used for making pudding).
146. *meloguey*(*hual*) 'rose' (most inclusive term).
147. *yueh.jih*(*hual*) 'month-season flower', *yueh.yuell horng* 'every-month turns-red'—i.e. '(monthly) rose'. Ch 488. Rd 454, BN 181 *Rosa indica*, L.
148. *chyangwei*(*hual*) 'hedgerose'. Ch 489. BN 1436 *Rosa acicularis*, Lindl.
149. *yee-chyangwei* 'wild hedgerose'. Rd 456, BN 987 *Rosa multiflora*, Th.
150. *jiuhual*, *jiuhual* 'chrysanthemum'. Rd 27, BN 1055 *Chrysanthemum sinense*, Sab.
151. *maantian-shing* lit. 'a skyful of stars' (a variety of chrysanthemum with clustered small yellow flowers). Ch 623. ≠ Ch 292, Rd 111, or BN 1251, 1252 under the same name, none of which the writer has any first-hand acquaintance with.
152. *kweihual*, *shianq.rhy-kwei* 'facing-sun *kwei*, i.e. sunflower, heliotrope'. BN 359 *Helioanthus annuus*, L. *kweihua-tzeel* 'sunflower seed'.
153. *shifan-lian* lit. 'western-barbarian lotus, *dahlia*-*hual*—i.e. 'dahlia'. Ch 662. BN 425 *Passiflora coerulea*, L.
154. *yuhlanhual*, (*shinyi* *L*) 'magnolia'. Rd 508, BN 281 *Magnolia conspicua*, Salisb.
155. *muhlan*(*hual*) Rd 'laurel magnolia'. Rd 510, BN 200 *Magnolia obovata*, Thunb.
156. *muhshiang-hual* lit. 'tree-fragrant flower'. Ch 491. Rd 453, BN 189 *Rosa banksiae*, R. Br.
157. *moh.li-hual* 'jasmin'. Ch 658. Rd 181 'Sambac or arabian jessamine', BN 698 *Jasminum sambac*, Ait., Lf 329 *Jasminum officinale* (which Rd regards as a different pl under No. 180).
158. *bair-lanhual* 'white *lan*-flower', so-called because of similarity of the blossoms to No. 141. It is a bush and has a much more aggressive scent. The flower (or pl) is called in the Wu dial. [bālehwo], without any suffix. The writer calls it *bair-lanhual* when speaking Mandarin, as do others who have occasion to call it. But the flower is little known in the north and therefore not much spoken of there. Unid.
159. *jiy.tzy.hual* 'gardenia'. Ch 728. Rd 82, BN 921 *Gardenia florida*, L.
160. *yingchuen*(*hual*) lit. 'welcome-the-spring (flower)'. Ch 350 Rd 179, BN 578 *Jasminum nudiflorum*, Lindl.
161. *ding.shiang*(*hual*) 'lilac'. Ch 613. BN 1095 *Syringa vulgaris*, L. ≠ 72.
162. *jin'yn-hual* 'gold-silver flower', (*reen-dong* 'endure winter' *L*)—i.e. 'honeysuckle' Rd 'or woodbine'. Rd 75 *Lonicera japonica*, Thunb.
163. *woanshiang-yuh* lit. 'evening-fragrant jade'. Ch 619. = 'tuberoze' (?) BN 180 (*yueh-shiah-shiang* lit. 'under-the-moon fragrant') *Polianthes tuberosa*, L.
164. *yeh.lai-shiang* lit. '(when) night comes (it is) fragrant'. Ch 660. = 'tuberoze' (?).
165. *shoel.shian*(*hual*) 'water-fairy flower, i.e. narcissus'. Rd 662, Lf 427 *Narcissus tazetta*, L., BN 209 var. *chinensis*, Roem.
166. *yuhtzan-hual* lit. 'jade-hairpin flower'. Rd 681, BN 279 *Hosta sieboldiana*, Engl.

167. *duh₂jiuan-hual* 'rhododendron' (?) Rd 201, BN 459 *Rhododendron indicum*, var. *macranthum*, Maxim.
168. *hwang-duh₂jiuan*, (*yang-buh-shyr-tsao* L lit. 'sheep-don't-eat grass') = 'azalea' (?) Rd 203, BN 408 *Rhododendron sinensis*, Sw.
169. *chlou-hae.tarng* 'autumn cherry-apple, i.e. begonia'. Ch 623a. BN 706 *Begonia evansiana*, Andr.
170. *ji₂guan-hual* Rd 'cockscomb'. Ch 343. Rd 559, BN 1478 *Celosia Cristata*, L.
171. *tzyy₂wei(hual)*. Ch 610. BN 1112 *Lagerstroemia indica*, L.
172. *huhtleel-hua(1)* lit. 'butterfly flower', (*liishyr* L)-i.e. '(Chinese) iris'. Rd 655 *Iris ensata*, Th., BN 1539 var. *chinensis*, Maxim.
173. *shiohchyou-hual* lit. 'embroidered-ball flower'. Ch 615. Unid.
174. *meeiren-jlau* lit. 'beautiful-woman *jiau*'. Ch 617. BN 718 *Musa coccinea*, Ardr.
175. *jiyishyang-tsao* lit. 'lucky-auspicious grass'. Ch 629. BN 354 *Reineckia carna*, Kth.
176. *jia₂jwu-taur* lit. 'between bamboo peach' (because the leaves look like bamboo leaves). Ch 656. BN 439 *Nerium odorum*, Soland.
177. *fenq.shian(hual)* 'phoenix-fairy flower', (*raan*)*jy.jia-tsao* '(dye) fingernail plant' (Rd *jyi-shing.tzy* 'quick-temperament')—i.e. 'balsam, touch-me-not'. Rd 296, BN 1308 *Impatiens balsamina*, L.
178. *shour.ji(sic!)₂hual*, (Ch *shukwei*, BN *yi-janq horng* 'ten-feet red')—i.e. 'hollyhock'. Ch 49. Rd 275, BN 1 & 1231 *Althaea rosea*, Cav.
179. *shiltzao-hual* 'take-bath flower', *lan.jy₂hual* '(face) rouge flower', *tzyy-moh.li* 'purple jasmin' (when it is purple), *juanq.yuan-horng* 'poet-laureate-red', *mow.li₂hual* (probably from *moh.li*)—i.e. 'four-o'clock'. Rd 242, BN 1102 *Mirabilis jalapa*, L.
180. *ternq₂luo*, *tzyy-ternq* 'wisteria'. Rd 418 *Wisteria chinensis*, DC., BN 1113 *Kraunhia floribunda*, Taub.
181. *charhua(1)* 'tea flower', (*shanchar* 'mountain tea' L)—i.e. 'camelia, tea oil tree'. Ch 768. Rd 266, BN 112 *The japonica*, Nois.
182. *gueyhual*, (*muhshi* L) Rd 'Cassia flowers', but American florists call it 'sweet olive'. Rd 184 *Osmanthus fragrans*, Lour.
- 182.1. *jinguey* 'golden guey', a deep yellow variety.
- 182.2. *ynguey* 'silver guey', a light yellow variety.
183. *hae₂tarng(hual)*, (*haehorng* L) 'cherry-apple flower' (pl or fl). Ch 684. Rd 437, BN 761 *Pirus spectabilis*, Ait. *hae₂tarng-guool* 'cherry-apple' (fr).
184. *meihual* 'plum flower'. Ch 697. Rd 447, BN 917 *Prunus mume*, S. et Z. *meijuang* lit. 'plum stump', alternate name used for the tree cultivated for the flower. *mei.tzy*, *ching-mei* lit. 'green plum', Rd 'dark plum'. ≠ No. 188.
185. *lahmei(hual)* 'wax plum', Rd also 'winter sweet'. Ch 679. Rd 504 *Chimonanthes fragrans*, Lindl., BN 1541 *Calycanthus praecox*, L.
186. *taur(1)shuh* 'peach tree'. Ch 698. Rd 448, BN 780 *Prunus persica*, S. et Z. var. *vulgaris*, Lf 539 *Amygdalus persica*, L. *taurhual* 'peach flower', name also applied to the tree cultivated for the blossoms; *taurl* (*taur.tzy* dial.) 'peach'; *taurhwul* 'peach stone'.
187. *shingshuh*, *shienqlshuh* 'apricot tree'. Ch 698a. Rd 444, Lf 539 *Prunus armeniaca*, BN 454 var. *Ansu*, Maxim. *shinghual* 'apricot flower', name also applied to the tree cultivated for the blossoms; *shienql*, (*shing.tzy* dial.) 'apricot'; *shingrel* '(Chinese) almond'. (Regular almond, Rd 443, Lf 405 *Prunus amygdalus*, Baill., is not so well known in China.)
188. *lii.tzy-shuh* 'plum tree'. Ch 704a. Rd 445, BN 453 *Prunus communis*, Huds. *lii.tzy*, (*jiaching-tzyy* dial.) 'plum'. ≠ No. 184. (The pulp of *mei.tzy* is thin, hard, and crisp, while that of *lii.tzy* is thicker and softer.)
189. *lishuh* '(Chinese) pear tree'. Ch 704. Rd 436, BN 922 *Pirus sinensis*, Lindl. *li F*, *ialli*, (*sheueli* 'snow pear' dial., *sheng-li* 'raw pear' dial.) '(Chinese) pear'. In the narrow sense *ialli* is a very juicy, crisp variety of *li*, cultivated in Shantung and neighboring provinces.
190. *lan.tair-li* lit. 'Cheefoo pear', *shi₂yang-li* lit. 'occidental pear'—i.e. 'pear'. BN 423 *Pirus communis*, L. Because of the length of the word, it is usually also applied to the tree without being compounded with the word *shuh* 'tree'.

191. *pynguoo-shuh*, *pyngguoo-shuh* 'apple tree'. Ch 705. Rd 434, BN 655 *Pirus malus*, L. *pynguoo*, *pyngguoo*, (*nay* L) 'apple'.

191.1-2. *hwu.li-bing* lit. 'fox-ice' or 'ice in the pot'; *huoo-lha-che* lit. 'fire-pulls-the-cart' or 'fire-drawn cart'. Small, crisp, sweet varieties of apples, grown in central and northern Hopei.

192. *shaguool-shuh* 'crabapple tree'. Ch 669. Rd 435 *Pirus malus*, L. var. *tomentosa*, Koch. *shaguool* lit. 'sandy (or granular) fruit', (*huahorn* lit. 'flower red' dial., *linchyn* L)—i.e. 'crabapple'. Rd equates this species with both apple and crabapple. Rd 437 uses the name 'crabapple' also for 183. In Japanese usage *riŋgo* (cognate with *linchyn*) is 'apple'.

193. *ing.taur-shuh* 'cherry tree'. Ch 699. Rd 449, BN 1528 *Prunus pseudo-cerasus*, Lindl. *inghual* 'cherry blossom'; *ing.taur* etymologically 'infant peach', i.e. 'cherry' AN -*ke* (2.101), because of the small size, also -*geh*.

194. *shyr.liou-shuh* 'pomegranate tree'. Ch 705. Rd 250, BN 375, Lf 276 *Punica granatum*, L. *shyr.liou* (*anshyrliou* L) 'pomegranate'.

195. *yangmei* Rd 'box myrtle' (pl or fr). Ch 674. Rd 621, BN 1182 *Myrica rubra*, S. et Z.

196. *shyh.tzy-shuh* 'persimmon tree'. Ch 701. Rd 188, CHPT 13.3, BN 645 *Diospyros kaki*, L. f. *shyh.tzy* 'persimmon' Rd also 'date-fig'. *shyhbiing*, *shyhbieengl* 'persimmon-cake, i.e. flattened dried persimmon'. *shyhshuang* 'persimmon-frost, i.e. whitish candy made of compressed sugar from persimmons'.

197. *tzaolshuh* 'jujube tree'. Ch 693. Rd *Zizyphus vulgaris*, Lam., BN 1029 var. *inermis*, Bge. *tzaol* 'jujube' Rd also 'Chinese date'. (*tzaol.tzy* dial., hence the custom of using the fruit as a propitious food at weddings and festivals, i.e. *tzaol-tzy* 'an early son'.)

197.1. *gar.ga-tzaol*, a northern variety, with rhombus, instead of the usual oval, vertical section.

198. *shan.ja* 'hawthorn' Rd also 'red haw'. Ch 706. Rd 422 *Crataegus cuneata*, S. et Z., BN 124 *Mespilus* c., S. et Z. *shan.ja(guool)* 'hawthorn berry'; *shan.ja-gau* 'hawthorn jelly'.

199. *tsaomei*, *yang-yangmei* lit. 'foreign box myrtle'—i.e. 'strawberry' (pl or fr). BN 510 (*herlan-mei* 'Dutch mei') *Fragaria virginiana*, Ehrh.

200. *shi.gua* 'western melon, i.e. watermelon' (pl or fr) Ch 678a. Rd 57, BN 419, *Citrullus vulgaris*, Schrad., Lf 438 (or *Cucurbita citrullus*) *guatzeel* 'watermelon seed'.

200.1. *sanbair-gua* lit. 'three-white melon', a variety with white pulp, white seeds, and whitish rind.

200.2. *maaling-gua* lit. 'horse-bell melon', a small, very sweet variety, common in southern Kiangsu.

201. *tyan.gua* lit. 'sweet melon', (*shiang-gua* lit. 'fragrant melon' dial.)—i.e. 'muskmelon' (pl or fr). Ch 676. Rd 58, BN 947 *Cucumis melo*, L. There are many varieties, all having thinner smoother skin and less fibry pulp than 'cantaloup' Bailey 955 *Cucumis melo cantalupensis*, Naud.

202. *pwu.taur* 'grape' (vine or fr). Ch 694. Rd 288, BN 1214, Lf 220 *Vitis vinifera*, L. AN for fr -*lih*, -*liell*, -*geh* (single grapes), AN for bunch -*guah*, -*du.lu*; *pwu.taur-terng(.tzy)* 'vine' AN -*ke*, -*jiah*; *pwu.taur-gal* 'raisin'. In spite of current orthography (Section 4), which makes it look like a binom, it was semanticized in very early times (see Lf 225) and is now still popularly analyzed as some kind of *taur* 'peach'. (Cf. Eng. *cranberry* as some kind of 'berry'.)

203. *lih.tzy-shuh* 'chestnut tree'. Ch 698b. Rd 610 *Castanea vulgaris*, Lam., BN 771 var. *japonica*, DC. *lih.tzy* 'chestnut'.

204. *her.taur-shuh* 'walnut tree'. Ch 669a. Rd 619, Lf 254 *Juglans regia*, L., BN 672 var. *sinensis*, Cas. *her.taur* lit. 'kernel peach', (*hwutaur* lit. 'foreign peach' L, [budo]⁸ Wu dial.)—i.e. 'walnut'.

205. *yee-her.taur* 'wild walnut, i.e. hickory'. *Carya*. Of the five species of *Carya* described in Bailey 325 the Chinese *yee-her.taur* seems to be nearest to *C. cordiformis*, Koch and *C. tomentosa*, Nutt.

⁸ Since the Wu dialect word for 'grape' is not [budo], as one would expect, but [bɔdo], no homophony is involved.

205. **jenshuh** 'hazel tree'. Ch 670. Rd 618, BN 1259 *Corylus heterophylla*, Fisch. *jen.tzy* 'hazelnut'.

(207. Rd *dushyan-tzyy* 'cashew nut'. Rd 311 *Anacardium occidentale*, L. The writer has not heard it spoken of in China.)

208. **feel.tzy-shuh** Rd 'kaya-nut tree'. Ch 705b. Rd 792, BN 1260 *Torreya nucifera*, S. et Z. *fei.tzy*, (*feishyr* L) Rd 'kaya nut'.

209. **baiguoo-shuh** 'ginkgo tree'. Ch 678. Rd 793, BN 153 *Ginkgo biloba*, L. *bair-guoo* lit. 'white fruit', (*ynshing* L lit. 'silver apricot', cognate of Japanese reading *gin'kō*)—i.e. 'ginkgo'.

210. **huasheng**, (*lawhuasheng* lit. '(when) fall the flowers, (nuts) grow' L, *charngsheng-guoo* lit. 'long-life fruit' dial.)—i.e. 'peanut' (pl or fr). Ch 686. BN 1197 *Arachis hypogaea*, L.

211. **ling.jeau** lit. 'ling-horn', (*ling*, *shoeli* lit. 'water chestnut' L)—i.e. 'water caltrop, horn chestnut'. Rd 243 CHPT 13.10, BN 558 *Trapa natans*, L. ≠ No. 48.

212. **bin(g).lang-shuh** 'betel palm'. Ch 702b. Rd 713, BN 1454 *Areca catechu*, Willd. *bin(g).lang* 'betel nut'.

213. **gan.je** 'sweet je, i.e. sugar cane'. Ch 703. Rd 756, BN 285, Lf 377 *Saccharum officinarum*, L. *tarng F* 'sugar'.

214. **pyi.ba-shuh** 'loquat tree'. Ch 702. Rd 427, BN 548 *Eriobotrya japonica*, Lindl. *pyi.ba* 'loquat'. The form 'loquat' is based on the Cantonese word, cognate with *lujyu* lit. 'reed orange' or 'marsh tangerine'.

215. **wuhua-guool** 'flowerless fruit, i.e. fig' (fr or pl). (In speaking of the pl the compound *wuhuaguoo(l)-shuh* is possible but usually avoided. There are two reasons for this. One is that of length. But a more important reason is that the immediate constituents of four-syllable compounds tend to be 2-2 rather than 3-1. Now, in the present case, a 2-2 analysis would make it 'flowerless fruit-tree', instead of 'tree of the flowerless fruit', as it should be. Hence the preference for either calling the pl by the fr or using a descriptive phrase *wuhua-guool de shuh*.) Ch 683. Rd 601, BN 1048, Lf 410. *Ficus carica*, L.

216. **jyu.tzy-shuh** 'tangerine tree'. Ch 695. Rd 347, BN 259, 1022 *Citrus nobilis*, Lour. *jyu.tzy* 'tangerine', *fwujyu* 'Fu(kien) tangerine', (Rd 'bitter peel tangerine').⁹

217. **mihgan-shuh** 'tangerine tree'. Ch 671. Rd 348 *Citrus nobilis*, Lour., BN 561 var. *mihgan* lit. 'honey-tangerine', (*gan F* dial.) 'tangerine'. (Rd 'sweet peel tangerine'.)

218. **chern.tzy-shuh** 'orange tree'. Ch 672. Rd 343, BN 731 *Citrus aurantium*, L. *chern.tzy*, (*goangjyu* lit. 'Kwangtung tangerine' dial.) 'orange'; *chernpyi* '(dried) orange peel' (used as medicine). The pron. *chern* is phonologically irregular, except that in the color name *cherng-hwang* 'orange' it ends in *-ng*.

219. **yow.tzy-shuh** 'pomelo tree'. Ch 695a. Rd 344 *Citrus decumata*, L., BN 652 *Citrus medica*, L. var. *acida*, Hook. *yow.tzy*, (*wendan* [vəntə], Wu dial.) 'pomelo' (like grapefruit, but drier and sweeter).

220. **ningmeng-shuh** 'lemon tree'. BN 516 *Citrus medica*, L. var. *limonum*, Hook. *ningmeng* 'lemon'. (The L form *yimuu-tzyy* lit. 'suitable-for-mother fruit' is a semanticization of the binom.)

221. **jinjyu-shuh** 'kumquat tree'. Ch 677a. Rd 345 *Citrus japonica*, Th., But BN 603 *Citrus nobilis*, Lour. var. *microcarpa*, Haask. *jinjyu* 'golden tangerine, i.e. kumquat'. Rd and BN add *geei-keh-cher(n)* 'give-guest orange' (because of the small size?). The English form 'kumquat' is simply the Cantonese pron. of *jinjyu*. Webster gives the literal meaning as 'golden orange', probably for the reasons stated in Note 9 to 216 above.

222. **shiangyuan-shuh** 'citron' (?) Rd 346a, BN 733 *Citrus medica*, L. var. *sacrodactylis*, Swingle. *shiang-yuan* 'citron' (?). The fruit the writer knows under this name is round, but the Latin variety name seems to put it into the next entry.

⁹ Since *jyu.tzy* 'tangerine' is much commoner than *chern.tzy* 'orange' in the Mandarin-speaking regions, the word has acquired the status of a name for 'the commonest citrous fruit'. Hence the practice, on the part of Mandarin-speaking persons in America, of calling oranges by the apparent misnomer *jyu.tzy*, since here it is oranges and not tangerines that play the part of the typical citrous fruit.

223. *for_oshoou-shuh*, (*gouyuan* *L* lit. 'hooked citron'). Ch 677. Rd 346 *Citrus medica*, *L.* var. *sacrodactylis*, Swingle, BN 429 var. *chirocarpus*, Lour. *for_oshoou* lit. 'Buddha's hand' (fr or pl).

224. *muhgua-shuh* lit. 'tree-melon tree'. Ch 701a. Rd 425, BN 1262 *Cydonia sinensis*, Thouin. ≠ 225. *muhgua* Rd 'Chinese quince'. This fruit is shaped like an elongated delicious apple. It has an oily yellow skin, with a strong fragrance, but is not good to eat.

225. *muhgua-shuh* lit. 'tree-melon tree, i.e. papaya tree'. Ch 688. BN 1073 *Carica papaya*, *L.* *muhgua* 'papaya'. (Ch *fangua* lit. 'foreign melon', but ≠ No 79, which is called *fangua* [vekwo] in Wu dial. Also ≠ 224.)

226. *mangguoo-shuh* 'mango tree'. Ch 682. Rd 312, BN 1378 *Mangifera indica*, *L.* *mangguoo* 'mango'. (BN also gives *mengguoo*.)

227. *lihjy-shuh* 'litchi'. Ch 672. Rd 300 *Litchi chinensis*, Sonn., BN 799 *Nephelium litchi*, Camb. *lihjy*, (*hijy*, *dambih L*) 'litchi nut'. Since this fruit is better known in dried form (with blackish pulp) than fresh (with white pulp), the fresh fruit is usually called *shin.shian lihjy* 'fresh litchi'.

228. *gueyyuan-shuh* 'longgan tree'. Ch 702a. Rd 302, BN 1414 *Nephelium longana*, Camb. *gueyyuan*, (*longyeen* 'dragon eye' dial.) 'longgan, lungngan' (the English form being the Cantonese pron. of *longyeen*).

229. *yangtaur-shuh*. Ch 684b. Rd 366, BN 146 *Averrhoa carambola*, *L.* *yangtaur*, (*wuiliann-tzy* *L* lit. 'five-converge fruit'), Rd 'Chinese gooseberry' or 'carambola'. It has a star-shaped cross section and an oval longitudinal section, which is narrower or wider according to the angle of section.

230. *chingguoo-shuh* 'Chinese olive tree'. Ch 674a. Rd 337, BN 1375 *Canarium album*, Raensch. *ching-guoo* lit. 'green fruit', (*gaan_olaan* dial.) 'Chinese olive'.

231. *yang-gaan.lan-shuh* '(foreign) olive tree'. Rd 183, BN 533, *Lf* 415 *Olea europaea*, *L.* *yang-gaan.lan* 'olive'.

232. *bajiau(shuh)* 'banana tree'. Rd 652, BN 286 *Musa sapientum*, *L.* *shiangjiau* 'fragrant jiau', i.e. banana' *AN -gel*. Note the difference between the names for pl and fr.

233. *ie.tzy-shuh* 'coconut tree'. Ch 675. Rd 720, BN 1177 *Cocos nucifera*, *L.* *ie.tzy* 'coconut'.

234. *boluomih, boluo* 'pineapple' (fr or pl). Ch 685 (*luhdou.tzy* lit. 'dew catcher'). BN 1310 (*fenqli* lit. 'phoenix pear') *Ananas sativus*, Lindl. ≠ Ch 684a, under the same name, with which the writer has no acquaintance.

235. *songshuh* 'pine tree'. Ch 714. Rd 789, BN 547 *Pinus*, *L.* (many species). *songguool*, *songtaa* 'pine cone'; *song.shiang* 'resin'. The Mandarin form *song*, in the first tone, does not agree with the ancient form *dz'iwong*, from which one would expect a second tone *tsorng* or *sorng*. In the central and southern dialects, which through Mandarin influence have assigned the usual character for 'pine' to the first-tone class, the regular (second-tone) form has continued to exist in everyday speech. Hence the occurrence of alternate characters for *tsorng* (see Section 4), sometimes with a slight variation in scope of application, for example to cover some spruces.

235.1. *maawoel-song*, lit. 'horse-tail pine'. Rd 789A.1. *Pinus masoniana*, Lambt.

235.2. *bair(pyl)-song* lit. 'white(bark) pine'. Rd 789B, BN 301 *Pinus bungeana*, Zucc.

236. *baeshuh, beanbor, beanbae* 'arbor vitae'. Ch 713. Rd 791, BN 869 *Thuja orientalis*, *L.* Also BN 641 *Chamaecyparis obtusa*, S. et Z., which Rd equates with 'juniper', but the illustration in BN is clearly the *baeshuh* as the writer knows it.

237. *baotaa-song* 'pagoda pine, i.e. spruce' Ch 713 (*tsyhbor*). *Picea*. Spruces are less common in China than pines.

238. *hwaishuh* Rd 'Chinese yellow berry, pagoda tree'. Ch 720. Rd 410, BN 1262 *Sophora japonica*, *L.*

239. *yang-hwai* 'foreign *hwai*, i.e. ash'. *Fraxinus*.

240. *yushuh* '(English) elm.' Ch 721. Rd 620 *Ulmus campestris*, *L.*, BN 1185 var. *laevis*, Planch. *yu.chyal* 'elm seed'. In the north it is boiled and eaten.

241. *yangshuh, bair-yang* 'poplar'. Ch 751a. Rd 623 *Populus tremula*, *L.*, BN 312 records some differences in species identification.

242. *leoushuh*, *yangleou* 'willow'. Ch 730. Rd 624, BN 655 *Salix babylonica*, L.
- 242.1. *chwei-yangleou* 'dangling willow, i.e. weeping willow'.
243. *fengshuh*, *horng-yeh* 'red leaf, i.e. maple tree'. Ch 750. Rd 463, BN 1184 *Liquidambar formosana*, Hce. Also BN 1323 (*chishuh*) *Acer palmatum*, Thunb. (?)
244. *ujlou(shuh)* 'black-mortar tree, i.e. tallow tree'. Ch 753. Rd 332, BN 790 *Sapium sebiferum*, Roxb.
245. *dongching-shuh* lit. 'winter-green tree', *donqching-shuh* lit. 'frozen-green tree'—Rd = 'evergreen'. Ch 766. Rd 310, BN 255 *Ilex pedunculosa*, Miq.
246. *wu₀torng-shuh*, (*chenn* L) 'plane tree', Rd also 'kolanut'. Ch 767. Rd 272, BN 923 *Sterculia plantanifolia*, L. F.
247. *rongshuh* 'banyan tree'. Ch 798. BN 1266 *Ficus wightiana*, Wall. var. *japonica*, Miq.
248. *shiang.chuen(shuh)*, *chuen₀shuh* Rd 'fragrant cedar'. Ch 750a. Rd 334, BN 1177 *Cedrela chinensis*, A. *shiang.chuen(tourl)* '(the edible) buds'.
249. *chow.chuen(shuh)*, (*chu, shu* L) Rd 'stinking cedar'. Ch 751. Rd 341, BN 1322 *Ailanthus glandulosa*, Desf.
250. *rowguey*, (*guey* L) '(Saigon) cinnamon' (pl or bark). Ch 715. Rd 495c, BN 414 *Cinnamomum loureiri*, Nees., Lf 541 *C. zeylanicum*.
- (251. *row-dowkow* Rd 'nutmeg'. Rd 503, BN 413 *Myristica fragrans*, Houtt. ≠ 108. The writer has not seen this plant.)
252. *wuujia-pyl* lit. 'five-add-bark'. Ch 722. Rd 234, BN 143 *Acanthopanax spinosum*, Miq. *wuujia_{pyi}(jeou)* 'w. liquor'.
253. *sangshuh* 'mulberry tree'. Ch 725a. Rd 605, BN 768 *Morus alba*, L. *sangyeh* 'mulberry leaf'; *sangrell*, (*sangshenn* L) 'mulberry'.
254. *feir.tzaw-shuh* 'soap bean tree'. (*feir.*)*tzaw-jia.tzy*, *tzawjeaul* 'soap bean'; *feir.tzaw* 'soap' (in the shape of balls made from the pounded pod), word also extended to manufactured soap, esp. in Wu dial. [bizo]. Ch 725. There are apparently two plants under this name: (1) Rd 387 'soap bean tree' *Gleditsia sinensis*, Lam., BN 467 *G. japonica*, (2) Rd 393 'soap bean' *Gymnocladus chinensis*, Baill., BN 556 H. Bn. For distinguishing the two, the writer's part as linguistic and anthropological informant fails here.
255. *chishuh* 'lacquer tree', Rd 'varnish tree, Japanese lacquer tree'. Rd 310, BN 1255 *Rhus varnifera*, DC. *chi F* 'lacquer' (from the sap).
256. *torng(you)shuh*, *youtorng* 'tung oil tree'. Ch 758 (*ingtzyy-torng*). Rd 321 *Aleurites fordii*, Hemsl., BN 1522 *Aleurites cordata*, Muell., Arg. *torngyou* 'tung oil'. ≠ 257, also ≠ 246.
- (257. *torng*. Ch 729. Rd 103, Lf 339 *Paulownia imperialis*, S. et Z. BN 784 *P. tomentosa*, H. Bn. The writer is not acquainted with this tree.)
258. *shianqpyi-shuh* 'rubber tree'. BN 349 *Ficus elastica*, L. *shianqpyi* 'rubber'; *shuhjiau* lit. 'tree-glue', i.e. 'raw rubber', less frequently '(prepared) rubber'. From the characters for *shianq* and *pyi* it would seem that they would mean literally 'oak bark' (see 265). But the 'immediate constituents' for the characters are not (1) *shianq* with radical 75 and (2) *pyi*, but (1) radical 75 and (2) the compound *shianqpyi* 'elephant-hide,' resulting in a complex compound of characters meaning 'elephant-hide of a vegetable nature'. With many binoms such a radical would be repeated for each character.
259. *tzongshuh*, (*tzongliu* L) 'palm tree'. Ch 764. Rd 719 *Chamaerops excelsa*, Th., BN 1189 *Trachycarpus excelsa*, Wendl. *tzong F* 'fiber' (from the bark).
260. *hwangyang(muh)* 'box tree'. Ch 767a. Rd 320 *Buxus sempervirens*, L. BN 1144 var. *mirrophylla*, Hk. F.
261. *jangshuh* 'camphor tree'. Ch 734. Rd 492, BN 1327 *Cinnamomum camphora*, Nees. *chaur.nao*, *jang.nao* 'camphor'.
262. *tarnshiang(shuh)* 'sandalwood', Rd 'red sanders wood'. Ch 735. Rd 590, BN 1426 *Santalum album*, L. *tarnshiang(muh)* 'sandal (splints)', used for incense.
263. *tzzy-tarn* 'purple sandalwood' (pl or wood). Rd 404, BN 1111 *Pterocarpus santalinus*, L. f. Its very heavy wood is used for making fine furniture.
264. *nan₀muh* 'nanmu tree'. Rd 502, BN 1185 *Machilus nanmu*, Hemsl. (several species used.)

265. *lihshuh*, *shianqshuh* '(Chinese) oak'. Ch 707. *lih_omuh* 'oak wood'. Rd 612, 612a, 613, 613a, BN 973, 1324, 649 *Quercus*, with several species and varieties.

266. *sha.muh*, (*shanmuh* L) Rd 'Japanese cedar'. Ch 733 (*shan*). Rd 786a, BN 454 *Cryptomeria japonica*, Don.

267. *sha.muh* Rd 'Chinese cedar'. Ch 734. Rd 786b, BN 447 *Cunninghamia sinensis*, R. Br. Ch describes in detail the difference between these two trees, which in popular usage, including that of the writer, are not clearly distinguished. Ch also notes that 266 *shan* is pronounced *sha* in the 'vulgar pronunciation' (*lii'in*) of the 'local dialects' (*tuuyeu*).

In concluding this account of Chinese plant words, the writer believes (1) that it accurately reports his own idiolect as far as plant words are concerned; (2) that his idiolect, apart from the exceptions noted, is a fairly representative sample of standard Mandarin; and (3) that the English and Latin equivalents give a good idea of what the things are that have been talked about. With (1) there is of course no room for argument.¹⁰ As for (2) and (3), the writer would be grateful for corrections from speakers of Mandarin and from sinologists and botanists.

¹⁰ Although one American reviewer, in *Harvard journal of Asiatic studies* 13.241 (1950), found my speech 'almost slangishly idiomatic' because I wrote out all the suffixes and particles and interjections used when I and my fellow Chinese professors of Chinese talk Chinese in China. I rather find that a compliment. Linguists will of course recognize here the familiar story that exact records of speech which sounds normal to the ear always look incredible and discreditable to the eye.

4. CHARACTERS

0.3. 梨, 杏, 李, 梨, 杏兒, 李子, *李, *李兒, 杏子, 榛子, 蓮子³, 松子兒³. 0.4. 紅葉, 脂麻, 麻油(香油), 脂麻醬(麻將), 桑葉, 荷葉, 榆樹的葉子, 樹林子, 個, 花草樹木, 植物, 一瓣花瓣兒, 根兒根兒根兒.

1.1. 稻, 薑, 蔥, 蒜, 參, 艾, 米, 筍, 藕, 漆, 菜, 草, 藥, 藤, 樹, 根, 顏, 皮, 油. 1.3. 兒, 子, 桃兒(瓜, 朥兒), 花兒. 這是一棵桃兒(樹). 蓮子³, 瓜子兒³. 一棵麥子, 一粒兒麥子, 一棵李子(樹), 一個李子, 竹子. 1.4. 跑堂兒. 1.5. 瓦松, 香菜. 松樹. 大, 字, 王, 先生. 1.6. 防風, 防風的藥. 迎春, 迎春花兒. 花生, 落花生. 1.7. 紅, 黃, 青, 紫, 白, 黑, 藍, 綠. 大, 小, 香, 臭, 甜, 酸, 苦, 辣, 草, 木. 山, 水, 野, 洋, 西, 南, (東, 北), 春, 夏, 秋, 冬. 1.8. 豆, 菜, 瓜, 草, 花, 果, 頭, 樹, 木, 子³, 瓜果, 蘆葦, 花草樹木. 1.9. 梔子花兒, 棒子面兒. 嘎². 嘎棗兒. 瓢兒菜, 芡兒菜, 鴨兒梨. 樹葉子, 菊花兒, 榆錢兒. 瓜, 朥兒. 豆兒, 花兒, 果兒. 1.10. 洋-白菜, 馬尾-松. 指甲-草, 野-核桃. 葵花-子兒, 綠豆-芽兒. 花-瓜, 朥兒. 白-蘭花兒. 1.12. 迎春. 賽雪梨, 月, 月兒. 紅, 春不老, 萬年青. 冬蟲夏草, 茵. 鐵樹開花, 蜘蛛抱蛋, 仙人過橋, 王不留行. 蘿蔔賽雪梨, 魚生.

2.1. 草, 樹. 植物. 棵, 一棵草, 一棵松. 種³, 一種³菜. (2.2. 植物, 草. 2.3. 一草一木.) 2.4. 花草樹木. 2.5. 穀子. (2.6. 五穀) 2.7. 雜糧. 2.8. 豆兒, 豆子. 2.9. 菜. 2.10. 葷菜. 2.11. 五葷. 2.12. 香料. 胡椒. 五香. 2.13. 素菜. (2.14. 菜蔬, 蔬菜.) 2.15. 草, 根兒, 樹,

木. 青草. 綠草. 2.16. 藥. 大. 黃是一種藥. 2.16.1 藥草. 2.17. 藤. 藤子. 2.18. 花兒, 朶(兒), 種一棵花兒. 2.19. 水果(~菓)(兒), 鮮果(~菓)(兒). 一棵杏兒, 一棵水果. 一個梨, 一根兒香蕉. 2.20. (水)果樹. (2.21. 灌木. 矮樹, 小樹, 一攢小樹.) 2.22. 樹. 2.22.1. 樹木. 樹林, 林子, 樹林子, 個, 片. 2.22.3. 森林, 片, 個.

2.23. 刺, 葉子. 2.24. 棵兒. 2.25. 根, 根兒, 根子, 個. 2.26. 鬚根兒. 2.27. 椿(子), 樹椿(子). 2.28. 頭, 菜頭, 蒜頭, 蔥頭. (2.29. 幹, 樹幹.) 2.30. 身子, 樹的身子. 2.31. 木. 頭, 塊. (2.32. 莖.) 2.33. 梗子. 2.34. 稈子, 稈兒. 2.35. 英子, 蘿蔔英子. 2.36. 杆, 旗杆; 桅杆; 杆兒, 欄杆兒; 桿兒, 桿子; 竿子, 竹竿兒, 箭竿兒. 2.37. 藤(子). 2.37.1. 藤絲(兒). 2.38. 樹枝, 枝子, 樹枝子, 樹枝兒, 枝兒. (桠杈, 樹桠杈, 桠枝, 樹桠枝.) 2.39. 葉子, 個, 張. 樹葉子, 樹葉兒, 葉兒. 2.40. 樹皮, 塊, 張. 2.41. 刺. 2.42. 瓢. 胚(~咕. 嘟)兒, 花瓢. 胚兒. (苞子) (藥, 女頭, 藥頭.) 2.43. 花兒, 朶, 種, 棵. 2.44. 花心兒, 攢花鬚. (雄蕊, 雄花鬚.) (雌蕊, 雌花鬚.) 2.45. 瓣兒, 花瓣兒. 個, 瓣, 一瓣花瓣兒. 單瓣(兒). 的. 雙瓣(兒). 的. 2.46. 花托兒. (萼) 2.47. 花粉. 2.48. 果兒(果子)(水果). (果, 實, 果實.) 2.49. 瓢兒. (肉(子)) 2.50. 皮. 皮兒. 2.51. 瓜. 2.52. 瓜果. 2.53. 蒂(~蒂)子. 2.54. 把兒⁴. 2.55. 乾兒, 葡萄乾兒. 2.56. 核(~櫟)兒. 2.57. 子兒, 個, 粒, 顆(≠棵!) 2.58. 種. 種子. 2.59. 殼兒. 2.60. 仁兒. 2.61. 穀子. 2.62. 穗兒. 2.63. 豆兒, 粒. 2.64. 豆角兒. (莢.) 2.65. 芽兒. 苗兒. 秧. (頭兒, 豌豆頭兒.) 2.66. 漿. 2.67. 汁(兒). 2.68. 油. 2.69. 棵兒, 梗兒, 葉兒, 瓣兒, 花兒, 花朵兒.

271. 麥芽(兒). 272. 糠, 礮糠. 273. 麩子, 麩皮, 麥麩(子).
 274. 米. 米粒兒. 275. 秫. 稽. 276. (綠)豆芽兒, 招菜. 277. 豆芽兒,
 豆芽兒(菜). 豆. 腐. 豆(腐)漿. 278. 菜薹(油菜). 279. 筍, 蘆筍.
 竹筴(子), 根, 條, 片兒. 280. 藕. 蓮. 蓬. 蓮子³. 荷葉. 281. 瓜子兒.
 282. 糖. 283. 松針(兒). 松. 香. 松果兒, 松塔. 284. 芸. 香. 285. 柳
 條兒. 286. 榆錢兒. 287. 桑椹(~葇)兒. (桑椹(葇).) 288. 漆.
 289. 樹膠. 橡皮. 290. 楓(樹)糖.

290. 樹. 棵. 種, 攢. 一樹的梨花兒, 一棵梨樹. 291. 個. 292.
 種(種). 種, 宗. 293. 棵. 293.1. 架, 一架葡萄(藤) (294. 科.)
 295. 枝, 一枝花兒. 296. 節兒. 一節兒竹子. 297. 根(兒), 一根兒麥稈
 子. 298. 張, 一張荷葉. 299. 朵(兒), 一朵黃花兒, 一朵兒玫瑰. 2100. 粒
 (兒), 一粒蠶豆兒. 2101. 顆, 一顆米粒兒. 2102. 牙兒, 鞞兒, 一牙兒
 橘子, 一鞞兒百合. (2103. 瓢.) 2104. 攢, 一攢矮樹. 2105. 嘟
 . 嚕, 一嘟嚕葡萄.

2106. 是. 這是一棵玫瑰. 玫瑰是一種花兒. 了, 花蔫了. 沒
 沒有. 2107. 這樹長葉子了, 這樹的葉子長出來了. (他死了一個朋
 友. 他朋友死了) 2108. 活了. 2109. 長了. 人參長葉子不長?
 2110. 出芽兒. 長苗兒. 2111. 灌漿. 2112. 發青. 2113. 發芽兒. 2114.
 長葉子. 2115. 長瓢. 肚兒. 2116. 開花兒. 2117. 結'. 結'果子. 這樹結
 石榴不結? 結²果. 結'子兒³. 2118. 生, 瓜太生. 2119. 熟了. 瓢
 兒熟了沒有? 熟了. 2120. 蔫了. 敗了, 謝了. 落了. 落葉子, 葉子
 落了. 落葉. 2121. 枯了. 2122. 死了.

2123. 他買的菜嫩. 2124. 發. 發白. (發)紅, 黃, 綠, 青, 藍, 紫, 白,

黑,灰,光(光,發光),毛,亮,暗,老,脆,甜,酸,苦,辣,一發嫩。
 2.125.長的。這樹長的高(這樹長高了),(這樹高)。 2.126.發,長
 的大,小,長,短,高,低,矮,粗,細,厚,薄,嫩,老,結實(結實²),
 肥,盛,旺。

2.127.早稻還沒有插秧。給。花兒匠得³給這棵菊花兒打
 頭了。 2.128.種。栽。撒子兒³。播種³(子)。 2.129.澆,澆花兒,澆田,澆
 水。 2.130.插秧。 2.131.插枝。分枝。壓枝。 2.133.打頭。 2.134.打藥。
 2.135.軋肥料。 2.136.修。 2.137.採。招。摘。擇。收,打。

3.0.1. BN:植物學大辭典,孔慶萊等編。

Ch:植物名實圖考,吳其濬。

Rd:本草綱目。李時珍。

CHPT:救荒本草,朱橚著(周定王)。

(3.)1. 麥子,小麥。麥穗兒;麥麵,麥粉;麥子,麥皮,麥麥(子);麥筋;
 麥芽(兒)。 2. 稻。米;飯;粥;米粉。 秈(~粳)米。糯米。糯米飯。
 3. 大麥。麥芽(兒)。 4. 小。米(兒),(粟) 5. 小。米(兒),(粱)。 6. 黍(子)(黍)。
 (7. 稷。) 8. 高。粱(蜀黍,蜀秫)。秫。稽。高。粱稗子。 9. 老玉米,棒子。
 棒子面兒。玉蜀黍。 (10. 燕麥。麥皮,麥皮兒,麥粥。) 11. 蕎麥。蕎
 麥麵。 12. 稗子。 13. 芝(~脂)。麻(胡麻)。麻油,香油;芝麻醬。
 14. 薏仁(米)(薏苡)。

15. 大豆,毛豆(兒),黃豆(兒)。豆芽兒(菜);豆(腐)漿;豆腐;豆油;
 醬;醬油;豆豉。 16. 蠶豆兒。 17. 扁(~菹)豆。 18. 雲扁(~菹)豆。
 19. 豌豆(兒)。豌豆苗兒。 20. 綠豆(兒)。豆芽兒,綠豆芽兒,招菜。
 21. 豇豆(兒)。 22. 刀豆。 23. 紅豆(兒)(赤小豆)。

24. 白菜, 黃芽(兒)菜(菘). 25. 油菜, 青菜, 白菜. 菜油; (紫)菜薹. 小白菜. 26. 洋白菜, 包菜. 27. 菜花兒, 個(朵), 標. 28. 芥菜. 29. 春不老, 雪裏紅. (芥, 青芥, 刺芥.) 30. 菠¹²菜(菠薐). 31. 莧菜(白莧). (31.1. 紅莧, 雁來紅.) 32. 馬齒莧(馬齒莧). 33. 薤菜, 空心菜. 34. 薺菜. 謝音菜. 35. 馬蘭頭. 36. 芹菜(水蘩). 37. 藥芹, 洋芹菜. 38. 茼(~同)蒿(菜), 蓬蒿, 蒿子. 39. 瓢兒菜. 40. 搨音(苦)菜. 41. 萁(~莖)菜. 42. 盧筍, 龍鬚菜. 43. 西洋菜. 44. 紫菜. 45. 海帶. 46. 生菜.

47. 蒿苳, 蒿筍. 48. 茭白, 茭耳菜, 茭兒菜(菰). 49. 茭蘭(甘藍). 50. 萁薺(~勃薺)(鳧次, 烏芋). 51. 蘿蔔(萊菔). 1.1 英子. 紅1.1, 白1.1. 賽雪梨, 蘿蔔1.1.1. 52. 胡蘿蔔. 53. 百合. 百合花. 個, 瓣兒. 54. 白薯, 紅薯(山芋), (甘藷, 白菽). 55. 山藥. 56. 山藥蛋兒, 土豆兒(洋山芋, 馬鈴薯). 57. 芋頭. 58. 慈姑(河烏菰, 剪刀草, 燕尾草).

59. 辣椒, 口椒, 辣口椒. 秦椒, 青椒. 60. 洋辣椒, 洋口椒(獅頭番椒). 61. 花椒, 川椒, (秦椒, 蜀椒). 62. 胡椒. 胡椒(面兒). 63. 八角, 11 茴香. 64. 薑, 生1, 塊. 65. 蔥. 66. 洋蔥, (胡蔥). 67. 蒜, 大1(小1)芽兒. 1 苗兒. (68. 葫. 小蒜為蒜, 大蒜為葫.) 69. 韭. 菜. 韭黃(兒). 70. 茴香(懷香). 71. 香菜(莞(~蕮). 荳菜)(胡荽). 72. 丁香, (丁子香). 73. 茶樹, (茗, 苦槩)茶葉, 茶, 青茶, 紅茶. 74. 咖啡. 75. 香草, (白茅香), (茅香).

76. 冬瓜. 77. 瓠子. 78. 葫蘆(~壺. 盧). 79. 倭瓜, 南瓜, (西壺盧), (北瓜), 番瓜, 倭瓜子兒³. 80. 絲瓜. 絲瓜絡(子). 81. 菜瓜.

(越瓜). 82. 黃瓜, (胡瓜). 83. 節瓜. 84. 苦瓜. 85. 茄子, (落蘇).
 86. 雞頭米, (芡實). 87. 金針(菜), 黃花兒菜, (萱草). 88. 西紅柿(番
 茄). 89. 蘑菇, 鮮蘑, 新鮮蘑菇, (蘑菇蕈, 肉蕈). 口蘑. (張家)
 口. 90. 香蕈, 鮮蕈. 香蕈, 香菇, 冬菇. 草菇, 花菇. 91. 雞縱(菌).
 92. 木耳. 93. 石耳. 94. 人參, 參. 1鬚. 95. 黃耆. 96. 甘草. 97.
 白朮(蒼朮). 98. 防風. 99. 黃連. (100. 見腫消.) 101. 冬蟲夏草
 (菌). 102. 麻. 黃. 103. 艾, 蘄(.州)1. 104. 大. 黃. 105. 蓖(~莊)麻.
 蓖麻(子)油. 106. 半夏. 107. 當歸. 108. 豆蔻. 109. 沙仁(兒). 縮
 砂蔻. 110. 薄荷(~菰蘭, 婆蘭). 111. 藿香. 112. 嬰. 粟花, (嬰子粟)
 鴉片. 1. 煙, 大煙, 煙土, 土. 113. 芸. 香.

114. 大麻, 麻子. 麻子; 麻; 1 繩兒; 1 繩. 115. 苧麻. 116. 葛.
 117. 棉(~綿). 花. (草綿); 棉花子兒; 棉花油. 118. 白及(~芡).
 119. 藤. 黃. 120. 通. 草, (通脫木). (121. 通草, 萬年藤.) 122. 燈心草.
 燈草, (燈心). 123. 葦子, 蘆葦. 蘆柴; 蓆(子), 蘆蓆; 笛膜兒. (124. 箬
 山白竹. 糲子) 125. 藤子. 126. 竹子. 筍. 冬1, 青1.

127. 蚰. 蚰兒草. 128. 狗尾巴草, (莠). 129. 蒲公英. 130. 菖蒲. (白
 菖, 石菖蒲.) 131. 爬山虎(兒), 爬牆虎(兒), (常春藤). 132. (鉤吻, 毒
 根, 黃藤, 野葛.) 133. 鳳尾草. (蕨). 134. 含羞草. 135. 瓦松. 136.
 仙人掌. 球, 棒. 137. 萬年青. 138. 萍草, 浮萍, 水萍. 蘋草. 水草,
 (藻). 139. 青(泥)苔. (140. 王不留行.)

141. 蘭花兒. 建1, 素心1, 葷心1. 142. (蘭花雙葉草.) 外國
 蘭花兒. 143. 芍藥(花兒). 144. 牡丹(花兒). 145. 荷花兒, (蓮花,
 蓮藕). 藕; 荷葉; 蓮蓬; 蓮子³, (蓮心); 藕粉. 146. 玫瑰(花兒).

147. 月季(花兒), 月月兒紅. 148. 薔薇(花兒). 149. 野薔薇. 150. 菊花兒. 151. 滿天星. 152. 葵花兒, 向日葵. 葵花子兒³. 153. 西番蓮. 大理花兒. 154. 玉蘭花兒, (辛夷). 155. 木蘭(花兒). 156. 木香花兒. 157. 茉莉(花兒). 158. 白蘭花兒. 159. 梔(~卮)子. 花. 160. 迎春(花兒). 161. 丁香(花兒). 162. 金銀花兒, (忍冬). 163. 晚香玉. (月下香). 164. 夜來香. 165. 水仙(花兒). 166. 玉簪花兒. 167. 杜鵑(花兒). 168. 黃杜鵑, (羊不食草). 169. 秋海棠. 170. 雞冠花兒. 171. 紫微(花兒). 172. 蝴蝶兒³花(兒). (蠶實). 173. 繡球花兒. 174. 美人蕉. 175. 吉祥草. 176. 夾竹桃. 177. 鳳仙(花兒), (染)指甲草, (急性子). 178. 淑氣花兒, (蜀葵, 一丈紅). 179. 洗澡花兒, 臙脂花兒, 紫茉莉, 狀元紅, 口口花兒. 180. 藤蘿, 紫藤. 181. 茶花(兒), (山茶). 182. 桂花兒, (木樨). 金桂. 銀桂. 183. 海棠(花兒), (海紅). 海棠果兒. 184. 梅花兒. 梅椿. 梅子, 青梅. 185. 蠟梅(花兒).

186. 桃(兒)樹. 1花兒; 1兒, (1子); 1棚兒. 187. 杏(兒)樹. 1花兒; 1兒, (1子); 1仁兒. 188. 李子樹. 1子, (嘉慶子). 189. 梨樹. 1, 鴨兒梨, (雪1, 生1). 190. 煙臺梨, 西洋梨. 191. 蘋果樹. 11, (奈). 狐狸(~壺裏)冰; 火拉車. 192. 沙果兒樹. 111, (花紅, 林檎). 193. 櫻桃樹. 櫻花兒. 櫻桃, 顆, 個. 194. 石榴樹. 1.1, (安11). 195. 楊梅. 196. 柿子樹. 11. 197. 棗兒樹. 1兒. (1子, 早子³) 嘎². 嘎棗兒. 198. 山楂(~樅). 1.1(果兒); 1.1膏. 199. 草莓, 洋楊梅, (和蘭莓). 200. 西瓜. 瓜子兒³. 三白瓜. 馬鈴瓜. 201. 甜瓜, (香瓜). 202. 葡萄. 粒(兒), 個, 掛, 嘟. 嚕; 1.1藤(子); 1.1乾兒. 桃. 203. 栗子樹. 11. 204. 核桃樹. 1.1, (胡桃, 蒲杏桃).

205. 野核桃. 206. 榛樹. 1子. (207. 都咸子.) 208. 榧子樹. 11, (1實). 209. 白果樹. 11, (銀杏). 210. 花生(落11, 長生果). 211. 菱. 角, (菱, 水栗). 212. 檳(~旃). 榔樹. 1-1. 213. 甘蔗. 糖. 214. 枇杷樹. 1-1. 盧橘. 215. 無花果兒. 216. 橘子樹. 11, 福橘. 217. 蜜柑樹. 11, (柑). 218. 橙子樹. 11, (廣橘); 陳皮. 橙黃. 219. 柚子樹. 11, (文丹). 220. 檸檬樹. 11. (宜母子) 221. 金橘樹. 11. (給客橙) 222. 香櫟樹. 11. 223. 佛手樹, (枸櫞). 佛手. 224. 木瓜樹. 11. 225. 木瓜樹. 11. (番瓜) 226. 櫟(~芒)果樹. 11. (櫟果.) 227. 荔枝樹. 11, (離枝, 丹荔). 新鮮荔枝. 228. 桂圓樹. 11, (龍眼). 229. 楊(~陽, 羊)桃樹. 11, (五斂子). 230. 青果樹. 11, (橄欖). 231. 洋橄欖樹. 11-1. 232. 芭(~巴)蕉(樹). 香蕉, 根兒. 233. 椰子樹. 11. 234. 波羅蜜, 11. (露兜子). (鳳梨). 235. 松樹. 松果兒, 松塔; 松. 香. 叢, 樅. 馬尾松. 白(皮)松. 236. 柏³樹, 扁柏^{2,3}. 237. 寶塔松. (刺柏.) 238. 槐(~櫟)樹. 239. 洋槐. 240. 榆樹. 榆. 錢兒. 241. 楊樹, 白楊. 242. 柳樹, 楊柳. 垂楊柳. 243. 楓樹, 紅葉. (槭樹). 244. 烏白(樹). 245. 冬青樹, 凍青樹. 246. 梧. 桐樹, (櫟). 247. 榕樹. 248. 香. 椿(樹), 椿. 樹. 香. 椿(頭兒). 249. 臭椿(樹). (樗).
250. 肉桂. (251. 肉豆蔻.) 252. 五加皮. 五加皮(酒). 253. 桑樹. 桑葉; 桑椹(~葢)兒. 254. 肥皂樹. (肥.) 皂莢子, 皂角兒; 肥. 皂. 255. 漆樹. 漆. 256. 桐(油)樹, 油桐, (罌子桐). 桐油. (257. 桐.) 258. 橡皮樹. 橡皮; 樹膠. 259. 椶樹, (椶櫚). 椶. 260. 黃楊(木). 261. 樟樹. 潮. 腦, 樟腦. 262. 檀香樹. 檀香(木). 263. 紫檀. 264. 楠. 木. 265. 櫟樹, 橡樹. 櫟. 木. 266. 杉. 木. (杉). 267. 沙木. 杉(俚音)(土語).

STETSON'S MODEL AND THE 'SUPRA-SEGMENTAL PHONEMES'

W. F. TWADDELL

Brown University

INTRODUCTION

00. The following discussion grew out of my attempt to fulfill a request from the Editor of *LANGUAGE* to review the late R. H. Stetson's *Motor phonetics: A study of speech movements in action*, 2d edition (pp. xi, 212, with 122 figures in the text; Amsterdam: North-Holland Publishing Company [for Oberlin College], 1951), hereafter referred to as *Mot. Phon.*² The routine reviewing procedures did not work out. Much of Stetson's material seemed unlikely to interest readers of *LANGUAGE*; many matters which might interest them, and which Stetson's topic might seem to bear on, are either not discussed at all or are only obliquely hinted at in *Mot. Phon.*² Further, Stetson's expository style, with its idiosyncratic terminology, its peculiar structure of repetitions and unsignaled transitions, its unlabeled analogies, its mixture of polemic and report, requires a paraphrase before discussion.

01. This paper is less and more than a review, in that it discusses only parts of *Mot. Phon.*², but discusses those parts as bearing on certain linguistic matters which were not of principal concern to Stetson.

02. Stetson's material relates to the sub-laryngeal processes of speech. No one denies the existence of those processes, nor that they have been investigated and taught less than the laryngeal and supra-laryngeal processes. The reasons for this relative lack of emphasis are probably two: the greater difficulty of objective measurement; and the conventions of Latin orthography, which traditionally has recorded the distinctive forms chiefly by notating the laryngeally and supra-laryngeally determined distinctions.

03.1. The lack of attention to sub-laryngeal processes is hardly healthy. Consider an analogy. Linguists decide to study the phenomena of speech, exclusively by listening and optically observing the articulatory actions visible to the unaided eye, working with subjects wearing high collars. The articulatory model thus arrived at would discriminate various lip-action categories; but all other heard phonemes would have to be attributed to unspecified factors of 'fullness, sharpness, roughness, thinness, depth, width, fuzziness,' and so on. No doubt a phonemic analysis could be made, and distributional properties of the several phonemes could be stated. But no modern linguist would willingly be without a knowledge of the total laryngeal and supra-laryngeal mechanism.

03.2. Whether or not we concede a sovereign relevance to articulation in the categorizing of the phonemes of a language, a knowledge of articulation is regarded as part of the equipment of a linguist. Trager, in *The field of linguistics* (*SIL*, OP 1, 2-3), remarks: 'The linguist must have ... a body of knowledge about the physiological activities ... he must be provided with descriptions ... of the articulatory movements of ... the organs of speech.' Bloch's Postulates (*Lg.* 24.10)

assume that an observer can be trained to identify any part of an utterance in terms of articulations—and presumably that he should be so trained.

03.3. So far as the laryngeal and supra-laryngeal articulations are concerned, linguists are today so trained. We are the heirs, through several generations, of 19th-century measurements and specifications of such articulations. Whether or not we have ourselves measured larynx posture or velum tension or tongue positions, we were taught to hear acoustic events in terms of those specifications; and consequently, by empathy, learned association, and proprioception, we can hear and classify with an assurance and rapidity otherwise impossible. In the laryngeal and supra-laryngeal domain the trained linguist of today can translate an unfamiliar speech sound into an articulation and relate it to other articulations, and he can describe it intelligibly.

04.1. Latterly, increased attention has been given to certain sound-features of speech which are not exclusively functions of the laryngeal and supra-laryngeal articulations. The logistic procedures which were developed for determining 'phonemes' (in the old-fashioned sense) have been applied to distinctions—psycho-socially demonstrable—of stress, pitch, transition, and various combinations of these. The general soundness of the procedures, long checked and confirmed in terms of articulatory specifications, has encouraged investigators and convinced their colleagues of the general validity of an analysis of stress, pitch, and transition-and-contour phonemes. But the analogy presented in §03.1 above is not far-fetched, as applied to the recent study of these phenomena.

04.20. Meanwhile, for many years, Stetson was measuring and specifying articulations in the sub-laryngeal domain. Using the methods of Rousselot to check certain hypotheses of Sievers, and proceeding behavioristically, Stetson studied the activities which move air through the larynx and the upper channels. Geography, training, and personal history kept him from the informal, face-to-face, frank-but-friendly discussions so characteristic of and so fruitful for American linguistics of the past three decades. Hence his research was not continuously enriched by the unfolding of phonemic investigation, nor did his research contribute its due stimulation or guidance to the study of sub-laryngeal phonemics.

04.21. Stetson's findings have not been directly useful for phonemics, for his stimulus-cues to subjects were selected to study movements and modifications of movements as such, not as the articulatory correlates of linguistically distinctive signals. Irritated by linguists' neglect of his findings, he scolded linguists for their logistic procedures.

04.22. Meanwhile linguists, like young patricians living off the income of a trust fund set up by a great-grandfather and unable to sympathize with the inelegant labor of delving, carting, and smelting, were proceeding through and beyond the articulatory specifications of upper-tract articulations, and were elaborating a mathematic of similarity and complementary distribution.

05. The gap is large. No one article can bridge it—certainly not this one. The conclusions of this article, in any honest appraisal, must be called anticlimactic. Nevertheless, some day linguists must know the relevant articulatory facts about ALL the vocal organs. If the sub-laryngeal movements are irrelevant for speech,

they must be SHOWN to be irrelevant, as—for example—the movements of the velum are not irrelevant. If sub-laryngeal movements are relevant articulations, the knowledge of those movements and of that relevance should be a part of the equipment of 'the trained linguist'; and he should be able to hear and classify certain acoustic events, by empathy, learned association, and proprioception, in terms of movements of the abdomen-diaphragm and the intercostal musculatures.

06. This discussion is meant as a preliminary survey of some possibilities for further study—study by individuals or teams. Section 1 is a sketch of some linguistic phenomena which are at present not yet correlated with articulations, as those phenomena are described by several investigators. Section 2 is an anthology of citations from *Mot. Phon.*² Section 3 is an expanded paraphrase of the anthology, with implications. Section 4 is a sketch of a bridge between the linguistic phenomena and the Stetson model. Section 5 is an appraisal.

07. Throughout this discussion I use the notation of Trager and Smith in their *Outline of English structure* (SIL, Occasional papers, No. 3; 1951), the most recent comprehensive treatment of American English phonology available in the summer of 1953. Future extensions and revisions of the analysis and formulations presented in that treatment may bring changes in notational practice. But since this discussion is concerned more with the observational data than with a particular phonemic interpretation, the notation here used should be readily convertible by later investigators or by those who today arrive at a different analysis of the observed phonetic facts.

SECTION 1

10.0. The linguistic phenomena that interest us are those collectively referred to as 'supra-segmental' phonemes. The term is controversial, and so are some of the phenomena. But if we list some classes of the phenomena, we can at least isolate the controversial elements.

10.10. PITCH is not controversial. There are variations in pitch in speech, and some of them are phonemic. Nor is the mechanism of production mysterious: since people can sing and can recognize melodies, enough attention has been paid to vocal pitch phenomena to familiarize us with the basic facts. The larynx is the organ responsible for the production of pitch-differentiations; modifications of vocal-band tension and position produce different fundamental frequencies of the voice-buzz, and these differences are potentially phonemic.

10.11. This function of the larynx is practically independent of the other three principal sets of speech-producing organs, with only two trivial reservations: (1) there must be sufficient excess of thoracic air pressure to force air through the glottis and drive the vocal bands; (2) there must be sufficient egress for the supra-laryngeal air to maintain, overall, a lesser pressure than that below the larynx. Given these two conditions, which are effectively normal, the pitch-controlling activity of the larynx is an autonomous speech function.

10.12. Of course, the determination of the number and kinds of pitch phonemes is a linguistic procedure. All that concerns us here is the mechanism. Whether Linguist A finds five pitch phonemes and Linguist B six, whether C thinks high-rising contrasts with mid-rising and D thinks that mid-rising is an allophone of

high-rising—all agree that there is a quite specific mechanism which controls those phonemic or allophonic pitch phenomena.

10.13. The one domain in which pitch-control is dependent upon non-laryngeal activities is duration. If it is alleged that a pitch phenomenon is phonemic, it must be demonstrated that the sub- and supra-laryngeal action is such that the pitch phoneme is audible, in terms of a minimum duration essential for pitch perception and identification.

10.2. STRESS. There is no controversy as to the presence of stress phonemes in some languages. There is general agreement that there are phonemic distinctions which are described as 'differences in degree of stress'. But the mechanism, the controls of these differences are seldom discussed in linguistic treatises. Usually, the differences are asserted or exemplified merely in terms of a 'more-less' scale; and beyond the principle that the differences are relative rather than absolute, there is little elaboration. In the conventional linguistic treatment, there may or may not be agreement about the number of stress phonemes in a language; the possibility of several kinds of stress-control, or even the nature of a stress-control, is rarely bruited.

10.31. 'JUNCTURE'. The term is more than controversial, it approaches the status of a shibboleth: either a ritualistic response or a verbum pudendum. Indeed, linguists as informants occasionally yield utterances of the type: 'He believes in juncture.' Some classic descriptions of juncture resort to 'the totality of phonetic features' (Trager and Bloch, *Lg.* 17.225), 'the way in which sounds are joined together' (Bloch and Trager, *Outline of linguistic analysis* 35), a 'structure point' (Trager and Smith, *Outline of English structure* 39), 'manners of terminating or going from parts of utterances' (Trager and Smith, *id.* 46). A somewhat different formulation is Bloch's (*Lg.* 24.41): 'fictions created ad hoc to account for the difference between certain sets of phonetically different segments.'

10.32. My assumption in this paper is that there are linguistically relevant distinctions which the believers in juncture call different junctures. Whether those distinctions should be notated with various quasi-Roman-alphabet symbols (as contrasting segmental phonemes) or with pseudo-arithmetical symbols (as different junctural or paracletic phonemes) is at present more a matter of taste, convenience, and public relations than of science. But as long as lay hearers can identify differentially what the juncturists call different junctures, I propose that we try to specify what differentiates them. If we can find some common articulatory feature which accompanies all instances of an alleged juncture, we may say that it specifies that juncture and is to be notated by a 'juncture symbol'.

10.4. SYLLABLE. The name is old and familiar, and appears frequently in contemporary linguistic treatises, though rarely with unambiguous specification. It is chiefly helpful to define 'syllabic' as noun and adjective, 'syllabicity', 'monosyllable', etc. In Section 3 below I have avoided the word (except in direct quotations from Stetson, who uses it as indicating a speech segment associated with a chest pulse). The procedure of this paper will be to tentatively assume chest pulses, examine possible varieties of chest pulses, and then see whether

that mechanism is such as to account for certain observed linguistic differentiations.

11.1. Since Stetson's experimental data are predominantly derived from American English, it is justifiable to take the observed phenomena of pitch, stress, transition, and syllabicity in American English as a starting point, and then examine the Stetson model, asking: 'Can this postulated mechanism produce these observed distinctions?'

11.2. First, let us consider some descriptions of some observed distinctions in the domains of pitch, stress, transition, and syllabic phenomena.

12.1. PITCH. There is general agreement that English syntax includes at least three and probably four different relative-pitch levels. (The difference between 'high' and 'extra-high' is conceded by all observers, but the linguistic status of the difference is sometimes disputed.) There is agreement, too, that there are numerous pitch allophones in the various stretches of speech corresponding to instances of various sequences ('231, 233, 232, 41') of the several levels, and to various lengths of such stretches—measured in segmental phonemes, or in syllables, or in terms of stress phenomena. Further, there is agreement that there are pitch factors in at least two different kinds of pre-pause terminals ('terminal junctures').

12.2. The pitch phenomena written ⁴ (extra-high), ³ (high), ² (average), ¹ (low) in the Trager-Smith notation are easily attributed to the laryngeal mechanism. In normal speech, the glottal tone fundamental corresponds to relative pitch. (The somewhat complex events involving pitch differentiation in whispered speech must be neglected for the time.) The laryngeal control of pitch implies that pitch-level phonemes and pitch-contour phonemes are associated with voiced segments. But the association of certain distinctive shifts in pitch with major stress, as in American English, is conditioned linguistically, not physiologically; it is a fact about American English (and other languages), not about the larynx. The speech mechanism, inherently, could be conditioned to produce pitch and stress differentiations quite independently of one another, or with a kind of association quite different from the American English kind.

12.3. The pitch factors in the transitions written (by Trager-Smith) as | (sustain), || (rise), and * (fade) are clearly controlled by laryngeal action; but other mechanisms are also involved. All observers agree that one acoustic component in the pre-pause termination /||/ is a rise in pitch beginning with the last pre-pause V and continuing through all following resonants. /||/ contrasts significantly with /*/ in that /||/ always has rise and /*/ never does. (/ */ contrasts significantly with /||/ in some features, but not significantly in pitch.) Beside the pitch difference between /||/ and /*/, there is a difference in the dynamic contour: typically, /*/ is described in terms of 'the voice dying away', 'fall[ing] off slowly', 'pre-pausal syllabics ... long or "drawled"'; whereas with /||/ 'the voice rises in pitch'. Which of these two factors—dynamic contour or pitch—should be regarded as primary, 'phonemic', and which as dependent? The question is linguistic; whatever the answer, the role of laryngeally controlled pitch differentiation can be specified in terms of coordination with the mechanism that controls dynamic contours. The only restriction is that if a pitch SHIFT

(e.g. a rise) is to be audible, it must be provided with a stretch of some duration during which at least a certain minimum loudness is maintained. The same requirement prevails if a pitch SUSTAIN (i.e. a phonemic non-shift) is distinct from a pitch shift.

12.4. After breath intake, in at least some pitch-and-stress patterns, there is a post-pause contour: a quick upward jump in pitch, which does not occur in close transition, nor after $/||/$ (sustain) or $/+ /$ (open transition). The musical notation of Paget-as-Shaw by Y. R. Chao (Transcribing reversed English, *Bull. Inst. History and Philology, Academia Sinica* 2.210-2 [1930]) shows some instances; and initial pitch jump is unmistakable in reversed playback of short utterances—assertions, commands, both kinds of questions. It is a contour allophone of pitches $/^3/$ and $/^2/$ at least; as such, it has the same phonological status as the contour allophones of $/^4, ^3, ^2, ^1/$ determined by $/||/$ sustain, $/||/$ rise, and $/\# /$ fade. Is the phoneme determining 'initial jump' an allophone of $/||/$ rise, or of $/\# /$ fade, or is it yet another phoneme? Is it an allophone of $/||/$ rise on grounds of direct phonetic similarity? Or is it an allophone of $/\# /$ fade, which has a falling-pitch variant, on grounds of mirror-image phonetic similarity—like $/y, w, h/$ in the Bloch-Trager-Smith notations? Harris, in *Methods in structural linguistics*, uses $/\# /$ for pre-utterance silence, describing it (15) as 'silence or break in the sequence of elements'.

13.1. STRESS. There is general agreement that there are three degrees of lexical stress in American English, which are customarily written $/', ^\circ, \sim/$, the convention being to leave $/\sim/$ unwritten. A variant of $/'/$, namely $/^\wedge/$, appears in connected speech, the distinction between $/^\wedge/$ and $/'/$ being grammatically significant.

13.2. Stress is usually defined in terms of relative loudness or 'intensity'. Thus Bloomfield, *Language* 90: 'stress which consists in speaking one of these syllables louder than the other or others'; Bloch and Trager, *Outline of linguistic analysis* 35: 'degrees of loudness'; id. 47: 'different grades or distributions of loudness'; Pike, *Phonemics* 63: 'increased in amplitude for greater stress'; id. 250: 'Stress: a degree of intensity upon some syllable which makes it more prominent or louder than an unstressed syllable'; Trager and Smith, *Outline of English structure* 35: 'different loudnesses ... consistent in their relative strengths'; Hall, *Leave your language alone* 75: 'different degrees of intensity of the air being expelled from the lungs: this intensity is termed stress'.

13.3. Discussion of the mechanism of stress phenomena is usually limited to explaining the degrees of loudness or intensity as resulting from more or less 'lung-action'. Thus Bloomfield, *Language* 110-1: 'Stress—that is, intensity or loudness—consists in greater amplitude of sound-waves, and is produced by means of more energetic movements, such as pumping more breath, bringing the vocal chords closer together for voicing, and using the muscles more vigorously for oral articulations.' Hockett, *Lg.* 18.5: 'the lungs are pushing.' Bloch and Trager, *Outline of linguistic analysis* 35: 'The degree of stress depends primarily on the force with which air is expelled from the lungs, secondarily on the energy with which the articulation is performed, on muscular tension, and on some other features—sometimes also in part on the pitch of the voice.' Pike, *Phonemics* 3:

'the walls of the lungs may contract so as to force air outward.' Bloch, *Lg.* 24.15: 'The aspect of stress is defined by the motion of the diaphragm and the other muscles that control the lungs.'

13.4. It seems fair to conclude from such quotations that the conventional linguistic explanation of stress differences is in terms of more or less vigorous expulsion of air from the lungs, with differences in the upper channel (larynx and above) corollary to the air-expulsion differences. Apparently the expelling musculature is usually conceived as one integrally functioning 'organ of speech'. We are thus invited to consider this organ as functioning with four distinctive degrees of vigor (for American English). In these terms, we could represent the stress phenomena by a line, with the left-to-right dimension depicting time and the up-and-down dimension depicting intensity. Or, if we were representing the articulatory process by an orchestral-score kind of notation, we could use one staff for stress, with notes placed on four arbitrarily selected lines or spaces; each greater degree of stress would be like-but-more-than a lesser degree.

13.50. However, the stress-producing mechanism, whatever it is, functions in some other linguistically relevant ways, over and above its role in producing four (or x) different loudnesses. The descriptions of 'junctures' suggests that time-intensity factors are involved in the four types of transition. These factors are contour-like, as contrasted with the point-like phenomena involved in the four stress phonemes, just as the pitch factors in the transitions are contour-like as contrasted with the point-like pitch phonemes.

13.51. Trager and Bloch, *Lg.* 17.225: 'Pre-pausal syllabics consisting of one of the so-called long vowels and diphthongs are exceptionally long or "drawled", the drawl extending over the whole syllabic; and the same is true of final nasals and liquids ending a stressed syllable. ... a loud stress on the last syllable before a pause falls off slowly and is accompanied by drawling of the segmental phonemes. A weak stress on the last syllable before a pause is usually still weaker than in other positions, and may decrease in loudness toward the end of the syllable. ... Final stops are usually unreleased; final voiceless stops are unaspirated; final voiced stops and spirants are partially unvoiced at the end; and final voiceless spirants are longer than at the beginning of an utterance. ... the features of open juncture are present not only before ... pause, but also internally in some utterances.' (Much the same description in Bloch and Trager, *Outline of linguistic analysis* 47.) *Structural notes and corpus* 4 (The Committee on the Language Program of the American Council of Learned Societies; Washington, 1952): 'Then the voice dies away, with no rise or fall of pitch, and there is a little pause before the next item. We indicate this by the symbol / # /. ... The voice, instead of dying away, rises slightly in pitch ... written with the symbol / // .' Id. 30: '... the sequence / (2) 3 ... / / ... The third pitch and the loudness of the voice continue to the end, and then there is a slight break. ... / // keep the pitch ... right up to the end of the phrase.' Id. 79: ' / (2) 3 1 | 1 ... // / ... The first ... has the familiar pitch sequence / (2) 3 1 /, but this time the voice continues instead of fading out. The second phrase ... has low pitch throughout, with a slight rise at the end.' Id. 10: 'In a phrase *with* a dividing point [i.e. / + /], the syllable before the dividing point is slightly drawled. In the forms that do not have a dividing

point, the first syllable is shorter and more hurried. In phrases with a dividing point, the syllable division is well marked. For example, it is before the /d/ in /déy driym/, but after the /d/ in /féyd àwt/. But in forms without a dividing point, like /gréydèyt/, the first syllable sounds as if it ends with /d/, and the second syllable sounds as if it begins with /d/, although there is only one /d/.
 Id. 15: 'Such a point of division ... may be symbolized ... as /+/. Sounds before /+/ are pronounced as if they were final, and sounds after /+/ ... as if they were initial.'

13. 521. The descriptions of the preceding paragraph may be summarized as follows, omitting references to pitch:

Pre-pausal sounds are allophonically long. The resonants are 'drawled'. The obstruents are prolonged. Voiced obstruents have a final portion without voice; voiceless stops are unreleased (i.e. eternal so far as that utterance is concerned) or unaspirated (i.e. released with minimum air-expulsion); voiceless fricatives are long.

Pre-pausal stresses are allophonically long: /'/' 'falls off slowly'; /~/ 'may decrease in loudness'. But pre-pausal /', ~/ (presumably also /^, '/) are further characterized by a decrescendo of intensity, a relatively gradual decrescendo, if 'dying away' can be read as implying a lingering demise rather than sudden death. This gradual decrescendo is compatible with 'drawling' of resonants and the relaxing articulation of obstruents described above. The feature of gradual decrescendo can be regarded as crucial, for it requires, as a corollary, some duration for its articulation and perception.

13.522. The two 'junctures' involved in gradual decrescendo are /#/ fade and /+/' open. The other two, /||/ rise and /|/ sustain, lack the 'dying away' feature; in them 'the voice, the loudness of the voice' 'continue to the end'. Allophonic length is a feature with /||, |/' also, for each has a pitch component (rise and sustension respectively) which requires duration. The loudness-contour of /||/ and /|/ is described as significantly level, ending with either an abrupt crescendo followed by cut-off, or with an abrupt cut-off, or with a decrescendo which differs from that of /#/ and /+/' by occurring after a conspicuously long level loudness.

13.6. Post-pausal allophones are described by Trager and Bloch, *Lg.* 17.225: 'Initial vowels may begin smoothly (with the glottis already in the position for voice) or with a glottal stop. Voiceless stops after open juncture are short, and aspirated even before a weak-stressed vowel; other consonants after open juncture are short. A loud stress on the first syllable after a pause sets in simultaneously with the beginning of the first segmental phoneme and rises rapidly in strength.' The post-pause jump in pitch (see §12.4 above) may have stress concomitants. As yet no observer has reported more than one kind of post-pause (including absolute utterance-initial) beginning, so far as allophones of /⁴, ³, ², ¹/ are involved. But there are indications that these point-like pitch and stress phonemes have characteristic post-pause contour allophones, and that these post-pause allophones differ objectively from utterance-interior and pre-pause allophones. The descriptions seem to point toward some kind of extra-strong onset which is quickly 'brought to normal level'—as though some basic driving

apparatus were being put into operation with a certain margin of excess effort, which is then adjusted to take care of actual requirements. Enough air is being expelled to drive an already adjusted glottis; voiceless stop closures, released, let out a notable quantum of air as aspiration; voiced obstruents are articulated to restrain the expelled air before the laryngeal obstruction is put into operation; consonants are short and rapidly give way to open-channel resonant articulations; loudness is conspicuous early in the transition from pause to speech; pitch begins 'too low', so to say, and vocal-band tension is rapidly adjusted to 'actual' (= phonemic average) pitch AFTER the beginning of audible utterance. If we may reverse the metaphor of 'the voice dies away', as applied to / $\#$ /, we may describe post-pause phenomena in American English as 'the voice is born prematurely'.

14.0. JUNCTURE. §12 and §13 have considered the pitch and stress phenomena which are described as being associated with /|/ sustain, /||/ rise, / $\#$ / fade, and, to some extent, /+ / open. The only other component of 'juncture' must be a matter of the relative timing of upper-channel articulations and variations in the rate of air-expulsion.

14.1. The impressively lengthened allophones of pre-pausal resonant sounds postulate a timing in which upper-channel articulations are unchanged or are changed slowly from one articulation to another, while the driving agent, the emitted air, continues in operation. What Stetson calls the final VC or VO of a breath group has a retardation feature, so far as upper-channel activity is concerned, and a persistence feature so far as air-emission is concerned; the retardation may be accompanied by a gradual stress decrescendo: / $\#$ /, or not: /|, ||/. It may be accompanied by a pre-inhalation posture: / $\#$, ||/, or not: /|/; by pitch rise: /||/, or not: / $\#$, |/.

14.20. The non-final allophones of phonemes determined by the upper channel are non-lengthened.

14.21. The post-pause allophones are not only short, they are crescendo (aspirated voiceless stops, voiceless-to-voiced 'voiced' obstruents, glottal stop as free variant of OV onset). The feature of rapid crescendo is so characteristic of post-pause allophones that in reversed speech the mirror-images ('final' in the reversed version) of initial sounds are frequently not perceived: such rapid decrescendo (reverse of rapid crescendo) and such brevity simply do not occur finally. In a phrase like *From an entirely different point of view*, the reversed version (on a tape loop) is transcribed by listeners without a final [f], whereas the original version is adjudged to have a wholly normal initial /f/. Perhaps all the features of 'initial' upper-channel phonemes should be regarded as allophones of a phoneme of 'preceding pause', or perhaps the initial features are determined by a positive, describable articulatory movement; the question is a linguistic one, to be answered only after all the phonetic data have been considered.

14.220. The distinctively different modes of transition within the utterance (i.e. not post-intake nor pre-intake) will be discussed at some length in §§37.40-472 below. Be it noted that there are three modes of transition, symbolized as unspaced letter sequence (close transition), /+ / (open transition), and /|/ (sustain). The principal difference between /+ / and /|/ is that /|/ is marked by

major phrase-final retardation and lengthening. The principal difference between 'close' and 'open' transition is that before open transition there is minor retardation and obstruents act as 'end of the syllable' (see §13.51 above).

14.221. Such statements as 'with /+/ the syllable division is well marked' demand a specification of syllable division which is not always forthcoming. Further, it is not clear whether syllable division is involved in the contrast between *school today* /skúwl+tədèy/ and *'s cool today* /s+kúwl+tədèy/; does the latter contain one more syllable than the former? And what of *If it's squarely in front* /f+ts+skwéhrlɪ | n+frént#/? Are there three syllable divisions (four syllables) in the stretch /ftsskwéhrlɪ/? Is there a syllable division (two syllables) in the stretch /nfrént/?

14.222. We will consider later (§15) the status of 'syllable' in contemporary phonemics. Here we need only note that two modes of transition within the breath group are described: /+/ and close transition. Leaving 'syllable division' as a criterion out of consideration, the characteristics of /+/ are described as minor 'final-like' retardation (prolongation) before /+/ and an 'initial-like' brevity and crescendo after /+/, whereas close transition is described as marked by 'normal' length and relatively gradual 'normal' progressive variations in stress and pitch. A sufficient but not necessary requirement for the occurrence of /+/ is /+^+^/ or /+^'/ or /'+^/ (see Trager and Smith, *Outline of English structure* 39); but /+/ occurs in other stress-pattern contexts also. The context /+^+, +^', '+^/ reminds us of Stetson's 'foot': see §§22.1-2 below: 'Every foot has its stress.' The final-like allophones before /+/ recall Stetson's 'the presence of an arresting consonant ... marks the foot division.'

15.0. SYLLABLE. Stetson uses the term as substantially equivalent to 'chest-pulse'. Pike (*Phonemics* 65) seems to accept this equivalence on the phonetic level: 'Segments are to phonemes as chest pulses (or phonetic syllables) are to phonemic syllables.' Pike characterizes his phonemic syllable (id. 246): '... one syllable represents a single unit of stress placement, or of tone placement, or of timing, or of vowel length, or of morphemic structure.' Bloch, *Lg.* 24.41, refers to 'the term *syllable*, which is often useful as the name of a structural unit determined by phonemic analysis'; but his discussion of postulates for phonemic analysis does not investigate the specification of syllables. Trager and Bloch, *Lg.* 17.223 ff., occasionally postulate a syllable without defining it: 'In the structure of the syllable vowels are nuclear, consonants marginal' (224); 'free syllables (*habitat, educate*) ... checked syllables (*habit, careful*)' (231); 'the intersyllabic consonant or cluster' (233); 'the consonant following the weak-stressed vowel goes with the second syllable (i.e. the onset of stress coincides with the beginning of the consonant)' (234). — Trager and Bloch's phonemic interpretation of a monophthong in *beat* as /iy/ (an interpretation which they have both; Bloch says, given up) must have implied some theory of syllable structure, whereby a syllable phase defines one of two phonetically identical segments as 'syllabic', the other as 'non-syllabic'; otherwise there would be a prohibited intersection of phonemes, since one allophone of /i/ resembles more closely (is identical with) one allophone of /y/ than either resembles its own co-allophones. This threatened intersection can be exorcised only by asserting that the /i/ and /y/

segments of the monophthong do not occur in the same position; and that assertion in turn implies that syllable-phase (syllabic vs. non-syllabic) is a determining element in the context. — In Bloch and Trager's *Outline of linguistic analysis*, 'syllable' is defined in terms of 'syllabic', and 'syllabics' are defined as 'sounds which constitute the peaks of sonority' (22): 'sonority' is either a primitive term or is equivalent to 'audibility'. In the discussion of semivowels (id. 22), we read 'a low vowel is more plainly audible than a higher vowel uttered with the same force'; 'two vowels may belong to the same syllable; either the first or the second may be more sonorous and act as the peak of the syllable; the other is said to be "non-syllabic"'; 'in many typical cases, the nonsyllabic vowel has a higher tongue position than the syllabic. ... Nonsyllabic vowels with a lower tongue position than a contiguous syllabic (but rendered less sonorous by receiving a weaker degree of stress) ...' (id. 23). This last qualification suggests that Bloch and Trager regarded sonority as the acoustic correlate of a resultant of (1) an 'openness' factor (id. 22: 'the sonority of a sound is determined primarily by the size of the resonance chamber through which the air stream flows') and (2) some factor of stress which is, however, not one of the stress phonemes. In Trager and Smith's *Outline of English structure*, 'non-syllabic' is introduced as a phonetic description (11); the final (tongue) position of the vowel nucleus in *seat* is marked by 'the symbol for non-syllabicity' (15). Thus 'non-syllabicity' seems here to be the primitive term. Throughout the *Outline of English structure*, 'syllable' and its derivatives are used as if in the expectation of being comprehended, and one key term, 'nucleus', seems to invite the interpretation 'nucleus of a syllable'. Stress is discussed as having a domain of some syllabic nature; a pitch change from lower to higher 'seems to actually take place at the beginning of the new syllable' (47-8).

15.10. Pike in his *Phonemics*, Hoenigswald in *JAOS* 64.155, and Eli Fischer-Jørgensen in *Acta linguistica* 7.15 ff. explicitly recognize a distinction between 'the phonetic syllable' and 'the phonemic syllable'.

15.11. Pike, *Phonemics* 65: 'I have concluded that phonetic and phonemic syllables differ.' Pike's wording implies an equivalence between 'chest pulse' and 'phonetic syllable'; this is probably to be taken as correlating some kind of speech-segment called 'phonetic syllable' with a kind of physiological event called 'chest pulse'. Pike's phonemic syllables are 'the pertinent structural units' (id. 144); they are correlated with placement of stress, tone, timing, duration, or even morphologically determined stretches (id. 246); 'in general, a phonemic syllable will be constituted of a single phonetic syllable with some rearrangement of the grouping in accordance with structural pressures' (id. 246). Pike's reasoning (id. 65) can be summarized: Let us assume that there are identical phonetic-syllable processes (chest-pulse activities) when an English-speaker says *lay* and a Spanish-speaker says *ley* 'law' and *let* 'I read'. But all three utterances will be described as monosyllabic by the English-speaker and as dissyllabic by the Spanish-speaker. Hence similar phonetic syllables may be different phonemic syllables in two different languages. — This conclusion would not be surprising, even if we accepted the assumption of some absolute 'phonetic syllable' identical in English and Spanish; but that assumption might well be questioned.

15.12. Hoenigswald, *JAOS* 64.155, specifically argues 'that phonemics can do without a phonetic definition of the syllable; that syllables are convenient or necessary terms for sequences of established phonemes.' He sets up instances of criteria for establishing (phonemic) syllables: (1) domain of accent phonemes, (2) minimum permitted sequences of (segmental) phonemes, (3) juncture phonemes determining syllable boundaries; these and like criteria will be of varying validity in various languages. Hoenigswald's phonemic syllable would be a feature of arrangement or order of phonemes which have been established and described without resort to phonetic-syllable considerations: an interesting undertaking, and possibly a desirable increase in rigor.

15.13. Fischer-Jørgensen, *Acta linguistica* 7.15 ff., takes a somewhat different view: the 'frame [for the definition of phoneme categories] must be some sort of phonemic "syllable".' Like Pike, she posits 'the phonetic syllable', but she is prepared for 'borderline cases from a phonetic point of view' and for 'syllable bases ... with different internal structure'.

15.140. The status of 'syllable' in contemporary phonemics seems somewhat unclear. Is it useful, is it necessary to determine and describe a phoneme of syllable boundary or syllable nucleus or both before other phonemes can be completely determined and described? If so, what are the procedures for establishing syllables? If not, why set up syllables at all?

15.141. The frequently encountered references to 'prominence, sonority, crests and troughs' leave something to be desired as criteria for phonemic assignments. No one, probably, would set up a phoneme /b/ on the grounds that its allophones sounded relatively cold or suave or sombre as compared with contrasting allophones of other, warmer or harsher or more lustrous phonemes: he would adduce labiality, voice, lenisness, or some other physiologically describable articulatory features, or he would specify some numerical acoustic constant(s). The attempt to establish syllable nuclei or boundaries in terms of 'relative loudness, audibility, sonority' is particularly tricky if discrete stress phonemes are also being established as significant degrees of relative loudness. Further, if it is desirable to have such units as are capable of instrumental specification, it should be remembered that the technical difficulties of recording and measuring acoustic 'sonority' are considerable.

15.142. Methodologically, it appears preferable to try to refer 'prominence, sonority, loudness' to some articulatory feature. If a component of muscular activity can be found which corresponds to a 'prominence' phoneme, it would have its place in the articulatory model of speech. Any comparable acoustic feature that could be found would fit into the acoustic model. For the present, it seems prudent to try to keep the two models distinct (expecting, of course, that important correlations exist between them) and not to operate with a hodge-podge of physiology (e.g. bilabial) and acoustics (e.g. sonorous, prominent).

SECTION 2

20. This section consists of direct quotations from Stetson's *Mot. Phon.*² These are excerpts, arranged in a sequence determined by the uses to which we

need to put his statements. The location of each excerpt in *Mot. Phon.*² is given in parentheses at its end; references are to pages, with a for the first column and b for the second.

21.1. 'Every utterance is a movement consisting of the phrase which is the larger, inclusive unit of which the breath group, foot, and syllable are organic parts. For this movement the posture is the adjustment of the abdomen-diaphragm and external chest muscles for regulated expiration; the movements of the breath group, and of the feet within the breath group are due to the abdomen-diaphragm movements of expiration. The series of small expiratory movements, the breath groups, constitute the phrase. The phrase is followed in the breathing cycle by a rapid inhalation. Although the utterance may consist of only one or two syllables, all the fundamental movement units are present.' (4b)

21.2. '... the breath group [is] due to an abdominal movement with its culminations which mark the stresses of the constituent feet. One of these stresses constitutes the main stress of the breath group, while the syllable pulses are produced by the intercostal muscles of the rib cage. ... After each movement of expiration for the breath group an intake may or may not occur; but there is always a readjustment to the slight change in volume of the chest. As the capacity of the chest decreases, the rib cage descends slightly and the abdominal muscles raise the diaphragm slightly to compensate for the outgo of air.' (3b)

21.3. 'The foot is the smallest unit group incorporating the syllables; it is due to an abdominal pulse which integrates a single stressed syllable or a few syllables grouped about a single stressed syllable.' (3b)

21.4. 'It is rather common but false to assume that the consonants and vowels float on the continuous pressure of the breath group. The chest is mistakenly thought of as the wind chest of an organ maintaining a steady flow under pressure throughout the breath group, interrupted by the constrictions (partial or complete) of the consonants, and colored by the shaping of the vowels. ... on the contrary, the chest does not maintain a steady pressure throughout the breath group. Instead, the chest muscles (intercostals) produce a separate pulse of pressure for each syllable; the pressure falls between the syllable pulses. The chest retains an overall posture which maintains its volume of air, but the minute contractions of the short muscles between the ribs (intercostals) force out little pulses of air which constitute the separate syllables. ... A slight pressure is generally maintained during the breath group, at least in English utterance, but the chest pulses of the syllables rise from this level. Between the breath groups the pressure goes to zero although the over-all posture and the chest volume are maintained; if there is an intake between breath groups, the pressure becomes negative.' (3ab)

22.0. Thus Stetson sets up a hierarchy of constituents: phrase (bounded by inhalations)—breath group (an abdominal movement having one main stress)—foot (an abdominal pulse with one stressed syllable)—syllable (a chest pulse). In detail, Stetson comments on the physiological activities corresponding to these units.

22.1. THE BREATH GROUP. '... the pulses of the feet and the breath groups which originate in the abdominal musculature.' (20a) 'The muscles of the

abdomen compress the viscera and force up the diaphragm, thus reducing the volume of the chest. The movement of the abdominal muscles makes the stress, grouping of the feet, and breath groups.' (193a) With diagram: 'Breath group musculature: R—The abdominal muscles led by the Rectus which exert pressure on the diaphragm through the liquid mass of the viscera. D—Diaphragm which opposes abdominal muscles. Muscle of inspiration. / These muscles act in opposition to produce the pulses of the foot and of the breath group.' (16b) 'The pressures from the abdominal muscles are transmitted through the viscera to the diaphragm. Thus the chest is maintained in position for the pulses of the intercostals.' (30b) 'The Rectus abdominis contracts in anticipation of the breath group involved; the syllables [nonsense syllables *ab-ba vak*] are assumed to occur no faster than three per sec., and an abdominal stroke (breath group) for each syllable is possible. If the syllable rate exceeds three per sec. the abdominal muscles fixate or make a slow, "tense", phrasing movement.' (57b) 'While the intercostals make a rapid series of syllable pulses, the abdominal muscles make a single movement against the opposing diaphragm for the group of syllables. This constitutes the breath group of one or more feet. The main stress of the group is produced by the culmination of this abdominal movement.' (30-1) 'These breath groups (of a single foot) may follow each other at the rate of 3-4 per sec.' (31ab) 'The breath group is an abdominal movement; the driving rectus and parietal muscles reciprocate with the diaphragm and thoracic muscles. This movement groups the syllable pulses and adjusts the chest-abdomen to the slight reduction in volume due to the outgo of the chest pulses.' (31-2) '... the syllables are parts of the breath group, and variations in the rate or the stress of the breath group may modify the component syllables radically.' (32a) 'The breath group may consist of one or more feet; the feet may consist of one or more syllables. Thus it is possible to have a breath group composed of a single syllable, or of 10-15 syllables.' (32a) 'Like the organization of the foot, the organization of the breath group is due to single movement of which the single syllables and the feet form a part, like the ripples on a larger wave.' (97b) '... the primary stress is the climax of the single slow movement which underlies and constitutes the unity of the foot, and of the breath group.' (97-8) 'The breath group is set off by its pause or brief intake, and is marked by the stress or stresses of the abdominal muscles which give the "culmination" to the breath group.' (162a)

22.2. THE FOOT. Reread the foregoing quotations, noting the references to 'foot' or 'feet'. Further: 'The feet are produced by the abdominal muscles which make the stress of the foot and group the syllables. The breath group is the movement of expiration involving the abdomen-chest adjustment which groups the feet in the breath group.' (57b) 'In the two-syllable feet, the iamb and the trochee, there are but two degrees of force. But the dactyl shows not only small variations in the duration of the three component syllables, but also gradations in the force of utterance. The stressed syllable is merely the climax of the stresses of the dactyl.' (97ab) 'Every foot has its stress, and every breath group has a dynamic climax as its primary stress.' (106b) 'Between little clusters of syllables, the feet, there may appear abutting or double consonants which link the syllables, or rather, link the closing and opening syllables of the feet composed

of syllables.' (162a) '... the presence of an arresting consonant, either alone or in an abutting (doubling) pair, marks the foot division within the breath group.'

(162ab) 'The foot is the rhythmic unit group composed of one stressed syllable only in the case of the one-syllable foot; composed of stressed and unstressed syllables in the case of trochees and dactyls, iambs and anapests ...' (195b)

22.30. THE SYLLABLE. With diagram: 'Chest pulse musculature: The External Intercostals which act to raise ribs in inspiration. The Internal Intercostals which act to lower the ribs in expiration. / These muscles act in opposition in the chest pulse.' (16b) 'The rapid pulses of the intercostal muscles for individual syllables ...' (16a) '... the intercostals make a rapid series of syllable pulses at rates as rapid as 8-10 per sec., and the larger muscles of the abdominal-thoracic musculature group and stress the syllables.' (31a) 'The internal intercostals working against the external intercostals make the syllable by a momentary pressure pulse. The internal intercostals draw the ribs down, deflating the chest; they are the muscles of expiration which produce the pulse. The external intercostals draw the ribs up, opposing the internal intercostals, and arresting the pulse.' (30b) 'The breath pulse from the chest is produced by the intercostals which do not adjust the chest for loss of volume. The expiratory movement which accommodates to loss of volume occurs in the abdomen.' (194b) 'The chest muscles are capable of producing 8-12 pulses per sec. This is the highest speed of small and well-coordinated muscles like those of the hand.' (35a)

22.31. 'The syllable may consist of an expiratory movement or chest pulse released and arrested by the intercostal muscles. This is the case with the single syllable consisting of a single vowel, OVO.' (33-4) 'If the syllable consists of a vowel only, OVO, the positive chest muscles (Inter. intercostals) start the wave of compression in the air column and the negative chest muscles (Exter. intercostals) stop the compression and take up the momentum of the chest wall.' (32b) 'These syllables, OVO, are ballistic movements with chest release and chest arrest. If the movement were not chest-arrested it would trail off in a sigh or a moan.' (34b)

22.32. 'If a consonant is incorporated, we may have the chest pulse released by the consonant, or the chest pulse arrested by the consonant. All phoneticians are agreed that the releasing and arresting consonants are unlike.' (34b)

22.33. 'If the syllable consists of a vowel and a consonant, OVC, the positive chest muscles start the wave of compression in the air column, the consonant stroke stops the compression wave, and the compressed air column helps to take up the momentum of the chest wall.' (32-3) 'In a series of syllables like *up, up* ..., the syllable pulse is started by the chest muscles; it is "chest-released". But the syllable pulse is stopped, arrested, by the consonant stroke. The consonant closes the vocal canal, and the rise of air pressure in the pharynx and chest takes up the momentum of the chest wall and brings the movement to a stop. ... the highest chest pressure of the pulse occurs when the consonant stroke closes the vocal canal. The lowest chest pressure between the syllable pulses occurs well after the consonant.' (45b)

22.34. 'When the syllable pulse is chest-arrested [OVO or CVO], the [external intercostal] muscles of the rib-cage act directly on the moving mass of the chest

walls to take up the momentum and arrest the syllable movement. But when the syllable pulse is consonant-arrested [OVC or CVC] the muscles of lips, tongue, velum, do not act directly on the mass of the chest walls. Instead, they close the vocal canal, and cause compression of the air column, which reacts on the moving mass of the chest walls. ... Often the syllable arrest is a combined consonant- and chest-action.' (51b) 'The consonants are auxiliary movements which release and arrest the syllable pulse, in lieu of a release or arrest by the intercostal muscles. Sometimes the auxiliary movement merely assists the chest release and arrest.' (159a) 'Every syllable has three invariable factors which (a) release the pulse either by the action of the chest muscles or by a consonant, (b) emit it by a movement which shapes the vocal canal for the vowel, and (c) arrest it either by the action of the chest muscles or by a consonant.' (200a)

22.35. 'If the syllable consists of a consonant and a vowel, CVO, the positive chest muscles [internal intercostals] start the wave of compression in the air column, which is momentarily blocked and then released by the consonant stroke, and the negative chest muscles [external intercostals] stop the compression and take up the momentum of the chest wall.' (33a) 'The releasing consonant stroke closes the vocal canal so that the action of the chest muscles compresses the air; then the rapid back stroke of the lip or tongue, with the opening of the jaw, releases the air. The beat stroke, and often the back stroke, occur during the beat stroke of the chest.' (44b) 'An auxiliary stroke may release the chest pulse, thus *ta, ta; tat, tat*. In snapping the fingers the movement of one finger releases the movement of the other.' (50-1)

22.41. SEQUENCES OF SYLLABLES AND SYLLABLE TRANSITIONS. 'When abutting consonants appear between syllables, the constriction is maintained throughout the pair; the first consonant of the pair arrests the chest pulse of the first syllable and the second consonant releases the chest pulse of the second syllable. A "double consonant" is a case in which the arresting consonant of the first chest pulse and the releasing consonant of the next chest pulse are the same consonant repeated.' (27b) 'The arresting consonant of one syllable and the releasing consonant of the next syllable may be linked.' (60a) 'In English, *topic* and *top pick*; *top egg* and *top peg*; *hit him*, *hit Tim*; *I do*, *I'd do*; *this eye* and *this sigh*; *unknown*, *un-own*; *I owe none*, *I own none*; *I'm Ike*, *I'm Mike*; *I lie*, *I'll lie* are phrases in which the contrast between the double and single consonant is marked. ... distinguish between *whole ode*, *hoe load*, and *whole load*, ...' (61a) 'The pressure for the "double" consonant is a bi-maximal curve, showing the arrest of the one syllable pulse and the back stroke of the chest muscles, before the increasing pressure for the release of the second syllable pulse.' (61b) 'The tracings of two different abutting consonants made by the same member, like *f-p* in *half-pay*, *m-b* in *humbug*, show precisely the same curve form as do the double consonants.' (65a) 'When series like *puf*, *puf* ... and *sat*, *sat* ... are uttered at increasing rate, abutting pairs of the form *f-p* and *t-s* result. Such abutting consonants are very much like double consonants.' (74b)

22.42. 'Two or more adjacent consonants may be classed as an intra-syllabic group when the group figures as a compound consonant in releasing or arresting the syllable movement, as contrasted with abutting consonants each of which

has a different function in two different syllables. ... In the group of the compound consonant the consonant movements are as nearly simultaneous as the nature of the movements combined will permit. Together they function as a single, arresting or releasing factor in the syllable.' (83a) 'These compound consonants are to be distinguished from abutting consonants. In the phrase *a tall D told E* the succession *l, d* occurs as an abutting pair in *tall D*, and as a compound arresting consonant in *told E* ... the releasing *d* of *l-d* contrasts with the arresting *ld* in the position of the maximum in the chest pressure in each case.' (85ab)

22.43. 'There is also the compound consonant in which the components maintain their identity and each has its own beat stroke, but the beat strokes occur so close together that they fuse with each other in arresting or releasing the syllable movement. ... a breath pulse does not occur between the strokes; there is little escape of air. (It is to be said, however, that ... such groups often break up into a preliminary silent syllable followed by a voiced syllable.)' (85-6) 'If the monosyllabic form is preserved, the two beat strokes must occur close enough together so that they are part of one releasing movement. Such compound consonants show a single form for the release of the chest pulse. But the pronunciation is often facilitated by breaking up the group into two abutting consonants; the first consonant becomes the arresting consonant of an adventitious syllable, and the second consonant releases the chest pulse of the original syllable.' (86ab) 'In the arresting form of the compound consonant, with an initial fricative like *s*, the mouth pressure curve is rounded if no second syllable develops. But an unvoiced second syllable often occurs, the form becomes bi-syllabic and the characteristic arresting-releasing form appears in the air pressure tracing.' (86b) 'The development of the bi-syllabic form is marked by the prolongation of the contact of the second occlusive, and by the arresting-releasing form of the air pressure curves which indicates the two chest pulses.' (87a)

22.5. RATE OF UTTERANCE AND SYLLABLE TRANSITIONS. '... the syllables are parts of the breath group, and variations in the rate or the stress of the breath group may modify the component syllables radically.' (32a) 'Sequences like CVO *pay, pay, ... tea, tea, ...*, however rapid, never show any interaction of the vowel with the following consonant.' (40-1) 'Sequences like OVC *ape, ape, ... eat, eat, ...*, show the arrest by the consonant and introduce an intersyllabic pause, if the syllables *ape* and *eat* are maintained; if the rate per sec. is increased beyond 3.5 per sec., the syllables become *pay, pay, ...* and *tea, tea, ...*, which are now chest arrested, while the arresting consonant has shifted to the releasing function in the next syllable.' (41ab) 'The chest arrest is faster than the consonant-arrest. The consonant mechanism involves a column of compressed air, while the muscles in chest-arrest act directly on the ribs of the chest. At a rapid rate the movements tend either to get into step or to drop in order to simplify the coordination; therefore the arresting consonant will drop while the releasing consonant retains its position because it comes in on the beat stroke of the syllable.' (71ab) '... as the rate of the syllable utterance increases, there is no time for the delaying process of consonant-arrest and the arrest must become a chest-arrest. This is the determining factor in the shift of the arresting consonant to the releasing position in the next syllable, or in the dropping of the arresting

consonant if the shift is not possible.' (51b)—'If the compound consonant occurs before or after a heavy stress at the division of the feet, it becomes two abutting consonants. "The Lor-dis my shepherd ...", "Thou prepareds-ta table ..."' (88a)

23.0. 'In discussing any system of skilled movements, there are three fundamental types of movement to consider: 1. The movement of fixation: opposing groups of muscles hold the member in position. When a person is about to speak, the chest is partly inflated and is often held in that position for a short time before any syllable is uttered. In such a case the chest is fixated. 2. The "controlled" or tense movement: ... at least two opposing groups of muscles work together in producing the movement. Both the antagonistic muscle-groups are contracted throughout the movement. The direction of the movement can be changed after it is under way; such a movement is relatively slow. ... The slow expiration of air in a prolonged vowel constitutes a "controlled" movement. The large breathing movement of the entire phrase is a slow, "controlled" movement. 3. The ballistic movement: the entire movement consists of a single pulse. ... The movement is started by a sudden contraction of the positive muscle-group which immediately relaxes. During at least half of the course of the movement neither of the antagonistic muscle-groups is contracted, so that the moving member flies free. At the end of its course the movement is usually arrested by the contraction of the negative muscle group. The movement is a movement by momentum. ... In speech, the rapid movements of articulation and the syllable pulse are ballistic.' (28ab)

23.1. 'It is possible to prolong a syllable at will, if the syllable is "open" or if the syllable ends with a voiced continuant like *m*, *n*, *l*, *r*, etc. ... the beat stroke of the syllable may be arrested by a back-stroke process which continues the movement in a slow, "controlled" form.' (36a) '... when the syllable movement has a consonant arrest, the length of the syllable is conditioned. The ballistic stroke of the syllable is arrested by the consonant movement and the syllable movement cannot be indefinitely prolonged into a controlled movement.' (58a) 'If the vowel is long in duration, the arrest of the syllable movement, and possibly the change to a controlled movement, will be well under way before the consonant is uttered.' (58b) 'The chest release and the chest arrest may or may not be phonemic.' (40b) 'The contrast ... between the "tense" vowel of the French *sité* and the "lax" vowel of the English *sit* is due to the fact that the French *sité* is arrested partly by the chest muscles, so that the *-te* may tend to become a separate syllable in emphatic speech. The English *sit*, on the contrary, is arrested by the consonant movement alone.' (42b) 'The syllable belongs to the type of movement characterized by a beat or pulse. Occasionally the movement may be prolonged when the ballistic pulse is continued into a prolonged, controlled (tense) movement.' (171b) 'There are cases in which the new chest pulse occurs without "diminution of intensity". Instead, a sudden rise in intensity marks the entrance of the new chest pulse.' (172a)

24.1 THE LARYNX. 'The larynx itself does not initiate the syllable nor control the process of articulation. ... The larynx makes possible the vocalization of speech. In just one thing is the larynx a prime mover: it determines the pitch of

the tone.' (30a) 'Glottal stop: when the consonant arrest is due to a sudden closure or stricture of the glottis.' (196a) 'It is only rarely, as in the "glottal stop" (coup de glotte) that the larynx acts as a true consonantal organ.' (37b) 'The glottis figures in two [consonant] sounds. The aspirate *h* ... is, strictly speaking, not a consonant. ... It is really a modification of the vowel and cannot occur in the arresting position. The rarer glottal stop figures as a consonant.' (49b)

24.2. 'It is not surprising that in speech changes of pitch should be noted at the stress; ... the heavy stroke of the accent involves the chest pressure and is apt to change the pitch because the laryngeal musculature is often affected by tensions in the other musculatures of speech.' (95b)

25. 'JUNCTURES AND RELATED UNITS'. 'Inter-clause (breathing cycle): the rapid intake indicates the sentence or stanza. Intra-clause, inter-phrase: the pause and intake mark the frontier between phrases. Intra-phrase, inter-breath group: the pause and possible intake mark the frontier between breath groups. Intra-breath group, inter-foot: the culminating stresses and the arresting consonant (abutting, double also) division of compound consonants mark the frontier between the feet. Intra-foot, inter-syllabic: shifts from arresting to releasing, fusion of consonants to compound consonants, various sandhi changes, mark the frontier between syllables. Intra-syllabic, inter-phonemic: shaping, releasing and arresting (delimiting) functions, movement fusions of compound consonants mark connections within the syllable.' (162-3)

SECTION 3

30.1. Section 2 presented a kind of first-level reorganization of some material from Stetson's *Mot. Phon.*², by excerpting and arranging according to certain tacitly assumed linguistic relevances. The words are Stetson's; but they are not all of Stetson, and their arrangement may be one he would have regarded as illogical or even misleading. We will not know, nor, for our purposes, is it permanently important to know.

30.2. In Section 3, I attempt to present a second-level reorganization of some of *Mot. Phon.*² Here the words are mine. I try to perform the acrobatic feat of paraphrasing and inferring from Stetson's work what I would have written (1) if I knew what Stetson knew about his research, as that knowledge is adumbrated in *Mot. Phon.*², (2) if I accepted that knowledge as being accurately determined and adequately established, and (3) if I was interested in certain linguistic applications of that knowledge (as Stetson was not). Only (3) is factually the case. I ask readers to join me in a momentary suspension of disbelief: let us, throughout Section 3, act as if the material in Section 2 is true; and let us see what kinds of speech mechanism would, under that heuristic assumption, be available for producing certain speech phenomena that we are interested in.

31. Stetson sets up four principal sets of speech-producing organs: (1) the abdomen-diaphragm, (2) the (internal and external) intercostals, (3) the larynx, (4) the supra-laryngeal organs. Of these, the larynx functions partly like the supra-laryngeal organs, in that its closure or stricture may release or arrest a chest pulse. Partly, the larynx functions independently, as the source of voice;

and voice can be regarded as something independent, something which is 'imposed' upon the other speech processes.

Thus, those basic speech processes involve the coordinated action of three sets of organs: the abdomen-diaphragm, the intercostals, and the larynx and supralaryngeal organs. These actions, and their coordinations, are learned behavior; none of them are simple respiratory acts; they are culturally conditioned, a part of language.

32.0 A continuum of speech can be unambiguously isolated by preceding and following pulmonic inhalations. (In this paper no account is taken of inhaled speech, implosives, or clicks; the phenomena here discussed are complicated enough when only exhaled speech is considered. The special problems of inhaled speech can well wait until we know much more than we now do about exhaled speech.) A segment between intakes is a 'phrase' or a 'breath group'; the phenomena at the beginnings and ends of phrases and of intake-marked breath groups are unquestionable terminal transitions. The physiological correlate of a breath group is a controlled movement of the abdomen-diaphragm musculature.

32.1. In this controlled movement, there is an overall dominance of abdominal action over the resistance of the diaphragm, so as to force the latter upward. This reduces the size of the chest cavity. Unless this reduction is compensated for by an escape of air from the lungs, that air will be under compression as a result of the abdomen-diaphragm action. To the extent that air escapes, the compression is less; if the compensation is complete, there will be no net increase of compression.

32.2. The rate of breath groups, as such, can be as rapid as 3-4 per sec. (*Mot. Phon.*² 31ab).

32.3. The unambiguous marker of breath groups is a preceding or following intake. But the positive features of a breath group (abdomen-diaphragm movement and any terminal features invariably accompanying intake) may characterize a speech segment as a breath group even without preceding and/or following intake.

33.0. The breath group as a whole is characterized by dominance of abdominal driving over diaphragm retarding. This is a controlled movement; but it need not be a uniform continuous movement. There may be, within the overall net upward displacement of the diaphragm, *SURGES* during which the diaphragm is driven with more-than-average intensity.

33.1. There are such surges. The coincidence of a surge with a chest pulse constitutes the culmination of a *FOOT*. If the breath group consists of a single foot, there is one and only one culmination. If there is more than one foot in the breath group, one of the culminations is the 'main stress'.

33.2. Thus there may be two degrees of intensity in the surges of the abdomen-diaphragm musculature: each such surge would represent a dominance of abdominal driving over diaphragm braking, and the control mechanism could be abdominal contraction and/or diaphragm relaxation (both relative).

33.3 If we need two degrees of abdomen-diaphragm surges to differentiate between the signal 'breath-group main stress' and the signal 'foot stress', the mechanism is thus available: extra intensity of abdominal action or momentary reduction of the diaphragm's opposing action or both. The nature of the muscles

involved in this 'controlled' movement would seem to permit such a more-or-less intensity of diaphragm lifting, resulting in two distinct degrees of compression of the thoracic cavity antecedent to intercostal activity. (I find no explicit discussion of this in *Mot. Phon.*², but Stetson's distinction between 'the stress of the foot' and 'the main stress of the [breath] group' seems to imply it. Hence, such two-level superimposed surges by the abdominal musculature acting against the diaphragm will be assumed in the following paraphrases.) Be it understood that this theoretical possibility of two kinds of abdomen-diaphragm surges is perhaps a maximum: there may well be languages which require only one degree of abdomen-diaphragm surge, occurring during a continued 'controlled' movement; indeed, some languages may operate without surges: a continuous, un-punctuated, 'controlled' upward displacement of the diaphragm would, in such a language, be sufficient in that it compensated overall for the loss of thoracic capacity resulting from chest-pulse expulsions.

33.4. In summary: the abdomen-diaphragm 'organ' displaces the diaphragm upwards, thus effectively furnishing a floor (rather literally) for chest-pulse action, so that any chest pulses must be compensated for by escape FROM ABOVE, through the upper channel. If the abdomen-diaphragm 'organ' has one or two degrees of surge, these are available for the chest-pulse mechanism: a chest pulse may coincide with one of the abdomen-diaphragm surges to result in a first- or second-degree intensity of air escape.

33.5. The abdomen-diaphragm activity can thus supply three kinds of signals (x is a surge; X is a major surge; is continuous upward displacement of the diaphragm):

[intake]	[intake]
[intake] x x x	[intake]
[intake] x X x	[intake]

(The number of x's is not significant; the position of X in relation to the x's is not significant.)

33.6. Since none of these activities is conditioned by respiratory requirements (the thoracic capacity could be reduced for breathing by relaxation of both abdominal and diaphragm musculatures), any one of these activities can be regarded as a linguistic activity, and, hence, as culturally determined.

34.0. Meanwhile, what of the chest-pulse mechanism?

34.1. The first observation is that the chest mechanism (the intercostals) is muscularly independent of the abdomen-diaphragm mechanism. Synchronization circuits may be, but need not be, set up between them; abdomen-diaphragm action may occur before, or with, or after chest action. Or (linguistically more relevant) there may be significant sequences of onset as between abdomen-diaphragm and chest pulses, or among different kinds of abdomen-diaphragm action and different kinds of chest pulses.

34.2. Given the (potential) compression of the thoracic cavity resulting from the abdomen-diaphragm movement, the escape of air could be either continuous or by pulses. The case of continuous expulsion might, of course, be regarded as a single pulse. In all other cases, there are two or more pulses. The mechanism which

controls the pulses of air escape is the rib musculature: the internal and external intercostals.

34.3. The internal intercostals contract. In and of itself, this action increases the compression of air in the thoracic cavity, by reducing that cavity. If air can escape through the upper respiratory channel, that escape compensates for the increased compression. If the channel is blocked by a barrier, the diaphragm would be forced downward, thus compensating (in part or whole) for the reduction of the thoracic cavity resulting from rib movement. But the diaphragm is being forced UPWARD (or at least braced against downward movement) by the abdomen-diaphragm movement. Hence the contraction of the internal intercostals produces a compression which is regularly relieved by escape through the upper respiratory channel.

34.5. (Clearly, the combined action of the abdomen-diaphragm set and the intercostals can be differently coordinated in time and in relative intensity. Since we are dealing here with learned behavior, such differences in coordination might be exploited for phonemic distinctions in a language; and it is quite possible that languages should differ in their exploitations.)

34.6. The contraction of the internal intercostals sets the chest wall in motion. This motion may be a 'free, ballistic' motion, if the internal intercostals contract, then relax, and there is no immediate opposing muscular action. Then the chest wall, once set in motion, will continue in motion until acted upon by some other force. The other forces which come into consideration are these: (1) some barrier to the escape of air through the upper respiratory channel, resulting in compression of air in the thoracic cavity, which in turn would brake and stop the motion of the chest wall; (2) a contraction of the external intercostals, which would act directly on the chest wall to brake and stop its motion; (3) a combination of (1) and (2); this combined action of the external intercostals and the upper-tract barrier(s) might be differently coordinated in time and relative intensity.

34.7. But the chest pulse need not be a 'free, ballistic' movement of the chest wall; it can be a 'controlled' movement, in which the internal intercostals and the external intercostals work against each other throughout the pulse, with net dominance of internal over external intercostal action until the end.

34.8. The rate of chest pulses, as such, can be as rapid as 8-10 per sec. (*Mot. Phon.*² 31a) or 8-12 per sec. (35a). (Recall that the maximum rate of breath groups is 3-4 per sec.) However, the high figure of 8-10 or 8-12 per sec. applies only to chest pulses which are braked and stopped ('arrested') by the external intercostals. A succession of chest pulses 'arrested' by an upper-channel barrier has a maximum rate of 3.5 per sec. (*Mot. Phon.*² 41ab).

35.1. The relations of timing of chest pulses and abdomen-diaphragm activities is not discussed in *Mot. Phon.*² Presumably Stetson's model permits the possibility of all combinations of chest-pulse varieties and abdomen-diaphragm surges or the continuous upward displacement of the diaphragm. Specifically, there could be as many kinds of 'foot stress' and 'breath-group stress' and 'non-stress' as there are kinds of chest pulse.

35.2. Nor does Stetson discuss the possibility of differentiating signals by various coordinations in time of various intensities of abdomen-diaphragm and

chest activities. Presumably only the abdomen-diaphragm stress-surges would be involved here. During such a surge, a chest pulse might begin relatively early or late in the sudden upward thrust of the diaphragm, with resultant difference in the total thoracic compression. Further, a diaphragm thrust accompanied by a 'free, ballistic' chest pulse might differ significantly from one accompanied by a 'controlled' chest movement. The free pulse might be briefer, and might be one degree less 'strong' on a phonemic scale, than the controlled pulse.

36. Stetson's material is almost exclusively from American English speakers. The few discussions of measurements of speakers of other languages (chiefly French) are based on the work of others using apparatus and aiming at objectives rather different from Stetson's. Hence it is necessary to reiterate that the speech activity of abdomen-diaphragm and the intercostals is learned behavior; it is entirely possible—indeed, it is likely—that there are considerable inter-language differences in this behavior. Just as an association between lip rounding and back tongue position is culturally conditioned, not physiologically determined, so various coordinations, sequences, and relative intensities of abdomen-diaphragm and intercostal action may be culturally conditioned. Such a statement as 'Language Q has only syllables of types CV, CCV, CCCV' may be another way of saying 'All chest pulses are consonant-released and chest-arrested' or 'Internal intercostal action is always followed by external intercostal action.' A statement 'Language R has no significant stress' might be paraphrased 'There is only one mode of abdomen-diaphragm action.' Language S might have a pattern of total sub-laryngeal activity such that chest action sets in simultaneously with abdomen contraction; but Language T might have a delay of 10–30 millisecc. Language V might have chest action continuing after abdomen contraction has ceased. Even the shape of curves which trace the thoracic compression resulting from intercostal contraction might differ from language to language.

37.0. Insofar as the larynx (as C-producing) and the supra-laryngeal organs are involved in the production and differentiation of the so-called segmental phonemes, their activity is the subject matter of classical phonemics. With that activity we are not here concerned, except insofar as allophonic variations are associated with sub-laryngeal activities.

37.10. Stetson sets up several kinds of chest-pulse action, taking account of the upper tract (some kinds of laryngeal and all supra-laryngeal activity). The variables are these: (1) chest or chest-and-consonant released; (2) chest or consonant or chest-and-consonant arrested; (3) 'free, ballistic' or 'controlled' movement. It should be noted that (1) above is conditioned by the 'type': OVO and OVC are necessarily chest-released, CVO and CVC necessarily consonant-released. But 'consonant-released' is not strictly comparable to 'consonant-arrested': at the release, intercostal activity is indispensable, with C momentarily restraining and then releasing the air compression produced by the internal intercostal action. But at the arrest, C can absorb the chest-wall momentum either *INSTEAD OF* or *IN CONJUNCTION WITH* external intercostal action. Hence, for (1) above, the linguistically relevant variable, in CVO and CVC chest pulses, is not the fact but the timing of the C constraint-and-release with respect to the Internal intercostal contraction. We may thus reclassify the variables: (1) early or late C re-

lease of the chest pulse; (2) fast (with chest arrest) or slow (without chest arrest); (3) 'normal' (free, ballistic) or 'prolonged' (controlled).

37.11. The releasing C of the first chest pulse in a breath group may be formed before or with the onset of thoracic compression resulting from a contraction of the internal intercostals; but it must be released after that onset, if it is to be audible. A priori we should expect the releasing C of a first chest pulse to have an allophonic form of relatively high audibility and distinctiveness, since it would have to be identified without the help of any significant allophonic effect upon preceding phonemes. Hence such a releasing C, when post-pausal, would be expected to be a later rather than an early release of the chest pulse, if the given language has a distinction between early and late C release.

37.12. An arresting C of the last chest pulse of a breath group may be formed with or without external intercostal action. A priori we should expect the relatively slower C arrest, as being more audible through its allophonic effect upon the preceding phoneme.

37.13. Those a-priori expectations are only that. Post-pause and pre-pause allophones are conditioned linguistically, not physiologically, though physiology may play a subsidiary role. The argument from audibility is more valid for voiceless than for voiced sounds, since voiced sounds have a more distinctive formant structure. And the factor of allophonic length vs. shortness is more relevant for stops than for continuants, since stops are more effective brakere-and-stoppers of the chest-wall momentum and therefore release or arrest the compression more rapidly. But with all reservations, it is worth our while to look for some sub-laryngeal components in post-pause and pre-pause allophones.

37.14. N.B. that a sequence of Cs may function as a post-pause releasing C or as a pre-pause arresting C, in the latter case with or without concurrent chest arrest. In the preceding and following discussion, C is to be understood as subsuming C, CC, CCC, CCCC, unless the contrary is explicitly stated, as in §§37.41-463.

37.2. In a discussion of post-pause releases and pre-pause arrests, the totality of sub-laryngeal activity—not merely the chest activity—may be relevant. It is possible that there is a characteristic synchronization or sequence of onset of abdomen-diaphragm and internal intercostal movements at the beginning of a breath group. And a controlled abdomen-diaphragm movement may end before, or with, or after, the arresting of the last chest pulse of a breath group. And the first or the last chest pulse of a breath group, or both, might be controlled rather than ballistic, or vice versa.

37.31. The first step is to define 'pause' with utmost rigor as an inspiration, a pulmonic intake resulting from action of diaphragm or external intercostals or both, with consequent enlargement of the thoracic capacity. So defined, 'pause' is a datum; and allophones identified in relation to it are indubitably post-pause or pre-pause.

37.32. The next step is to determine the categories of sub-laryngeal activity which accompany the unquestionable post-pause and pre-pause allophones. Sequence of onset of abdomen-diaphragm activity, internal intercostal chest pulse, formation-restraint-release of upper-channel C—these could be the de-

terminants of post-pause C allophones. *Mutatis mutandis*, likewise for pre-pause allophones. And for both, relative intensities of activities of the three musculatures might be relevant. The inquiry may be complicated by considerable variety of post-pause and pre-pause activity within the sub-laryngeal mechanism itself. The first chest pulse may accompany a MAXIMUM (breath group stress), a MAJOR (foot stress), or NO abdomen-diaphragm surge. If the language displays this kind of variety, there may be several sets of post-pause and pre-pause C allophones.

37.33. The next step is to discover whether these post-pause and pre-pause allophones are significantly different from what must be regarded as other allophones of the same C, if the phonemic identification of C is determined by only upper-channel activity, e.g. as post-pause [b] is phonemically classed with pre-pause [b] and with [b]s during a breath group, on the basis of articulations in and above the larynx. If there proves to be a constant difference (in terms of sub-laryngeal activity) between post-pause and pre-pause allophones of C and some other allophones of C, then it is worth while to dig deeper.

37.34. If all intra-pause C allophones differ considerably from all post-pause and pre-pause C allophones, then the pause is to be considered the determinant of the allophonic variation, and the intra-pause continuum is a relevant segment—certainly relevant phonologically, and probably also on the morphemic or syntactic level or on both.

37.35. However, if some but not all intra-pause C allophones differ considerably from all post-pause and pre-pause C allophones, then we must consider the sub-laryngeal components as themselves unpredictable in terms of pause (pause = pulmonic intake). If unpredictable, they are significant, i.e. phonemic. The phoneme corresponding to that bundle of components is then distributed so that it always occurs at pause and sometimes occurs elsewhere.

37.36. There may be a suspected phoneme of sub-laryngeal components which occurs at pause (= pulmonic intake) and elsewhere, but only in conjunction with, say, pitch shift, or a particular single sub-laryngeal component (controlled chest pulse, or maximum abdomen-diaphragm surge). Then that conjunct feature could be regarded as the phonemic determinant, and as occurring, if you will, at a non-intake allophone of pause. This interpretation would be confirmed if we found that these non-intake 'pauses' can be 'replaced with' intake pauses in various styles or by various speakers producing 'the same utterance'. In such cases, we might prefer to regard the articulatory bundle as marking a phoneme which occurs at pause, and regard pause as having a free variant, intake.

37.37. The number of sub-laryngeally determined phonemes in a given language is probably small (under a dozen, presumably). Possibly there are languages with no sub-laryngeal phonemes—languages in which the abdomen-diaphragm movement is an uninterrupted steady controlled movement to compensate for escape of air from the thoracic cavity, and in which all chest pulses are released and arrested in one mode. But there are other languages in which these musculatures operate with significant differentiations. One kind of differentiation we have described: a bundle of sub-laryngeal components, or a determining feature, which has unpredictable occurrence between intakes, and which conditions allophones

of upper-channel phonemes. There could be several such sub-laryngeal determinants, i.e. several sub-laryngeal phonemes.

37.38. We have so far considered the possibility of allophones at 'pause' (intake or non-intake) differing from other allophones. A plausible interpretation is that such allophonic variations involve distinctive activity of the abdomen-diaphragm either alone or in conjunction with the intercostals. Stetson's definitions of 'phrase' and 'breath group' (in §21.2 and §22 above) point toward a correlation between these segments and the abdomen-diaphragm movement.

37.40. Another kind of sub-laryngeal activity which may be phonemic has to do with the sequences of chest pulses and their immediate relations with Cs. Here we are specifically dealing with intra-pause phenomena.

37.41. What if we find, between two successive chest pulses within a breath group, one particular C segment, as determined by upper-channel activity? (See §37.33 above.) Is it phonemically distinctive whether that C functions to arrest the first pulse (if so, with or without concurrent chest arrest) or to release the second pulse (if so, early or late)? The answer may, of course, be different for different languages. Let us assume that for a particular language the answer is Yes.

37.42. The timing of chest pulses in relation to upper-channel activity is assumed to be distinctive. If the first chest pulse is still proceeding as a free ballistic movement when the C-obstruction is formed and is then arrested by that obstruction, the result would be a relatively slow arrest with major allophonic effect of C upon preceding V. This timing would favor an open transition (probably resembling pre-pause) between two successive chest pulses.

37.43. But if the chest pulse is already being arrested by action of the external intercostals when the C-obstruction is formed, the arrest would be relatively fast and the allophonic effect of C upon preceding V relatively less; and a closer transition between the two chest pulses would be favored. This closer transition is especially favored when the chest arrest of the first pulse and the chest release of the second pulse occur while the C-obstruction is in effect.

37.44. The timing of the second pulse in relation to a C-obstruction could be distinctive in terms of early or late release. If the second chest pulse is in its early stages when the C-obstruction is removed, close transition is favored; if late, open transition.

37.45. Stetson describes some 'adventitious syllables' (see §22.43 above): the occurrence of two pulses of the internal intercostals, one 'silent' and one 'voiced', during supra-laryngeal movements corresponding to a compound consonant. Stetson treats these two-pulse compound consonants as a kind of optional, 'facilitated' pronunciation of a one-pulse compound consonant. But given two timings of concurrent actions, the differences between them might be PHONEMIC: the 'same' sequence of consonants as ascertained by supra-laryngeal actions might be significantly different in their transition features; the one-pulse sequence would presumably have close transition, the two-pulse sequence an open transition.

37.460. We may represent some different possibilities of transition, using these symbols:

V, unobstructed escape of air through the upper channel
 I, internal intercostal action
 E, external intercostal action
 C, C-obstruction by upper-channel action
 Cf, formation of C-obstruction
 Cr, removal of C-obstruction
 ECfE, Cf during E
 ICrI, Cr during I

Then we might have such various types of transition between successive chest pulses as these:

- (1) VCfCrIV (Cf without E; Cr before I)
- (2) VCfICrV (Cf without E; I before Cr)
- (3) VECfICrV (E before Cf; I before Cr)
- (4) VECfEICrIV (Cf with E; Cr with I)

These are not all the possibilities; but they serve to illustrate the extent of variety. In conventional notation, these four types might correspond to American English:

- (1) VC + V (*bad apple*)
- (2) VC + CV (*could do*)
- (3) V + CV (*we devils*)
- (4) VCV (*widow, reduce*)

37.461. In §37.460 we were taking C as some one particular obstruent phoneme, whose formation and removal occur during different phases of two successive chest pulses. Much the same variety is possible with compound Cs (CC, CCC, CCCC). Thus, if a compound C consisted of /-rbl-/, for example, the various coordinations of intercostal and upper-channel activity could yield:

- (1) Vrb + IV (*curb-light*)
- (2) Vrb + blV (*verb-blest*)
- (3) Vr + blV (*far-blown*)
- (4) VrblV (*burbling*)

In terms of the narrative account in §37.460, Cf and Cr would here relate to the lip obstruction /b/ in the compound C /-rbl-/.

37.462. Similarly, if *Act 8* is /ækt + ért/, then the formula (1) VCfCrIV implies that the release of the /t/ C-obstruction precedes the internal intercostal action of the second chest pulse. And for an anti-aircraft gunner known as *Ack Tate* /æk + tért/, (2) VCfICrV implies that Cf relates to /k/ and ICr to /t/. In *actate* 'to officially put on the record' /æktèrt/, (4) VECfEICrIV implies that the /t/ obstruction is removed before I is completed.

37.463. The treatment of a venerable quartet would be indicated as follows:

- (1) *night-rate*: /t/ formed and removed before I
- (2) *night-trait*: /t/ formed without E, removed after I
- (3) *Nye-trait*: /t/ formed after E, removed after I
- (4) *nitrate*: /t/ formed with E, removed during I

37.470. All the discriminations in §37.460 involve some C-obstruction in relation to two successive chest pulses. If there is no C sufficiently obstructive to compress the air column and thus arrest the momentum of the chest wall, that arresting may be performed by activity of the external intercostals: a chest arrest. Stetson's data do not help in classifying different possibilities of coordination of chest and upper-channel activities in this case. But the scheme of §37.460 seems adequate to take care of such stylistic or regional variations as *chair* /čEYə, čEY + yə, čEH + yə/. Likewise various transition forms in *go on, be a*.

37.471. Presumably the differentiating mechanism would be much the same as for the transitions described in §37.460, with 'S, semivowel action' replacing 'C, consonant obstruction'. Thus we could set up: (1) VS + V: VSfESrIV (semivowel formed before braking by the external intercostals, released before action by the internal intercostals: /gow + ɔhn/); (2) VS + SV: VSfEISrV (semivowel formed before braking by the external intercostals, released after action by the internal intercostals: /gow + wɔhn, bɪy + yə, čEY + yə/); (3) V + SV: VESfISrV (semivowel formed after braking by external intercostals, released after action by internal intercostals: /goH + wɔhn, čEH + yə/); (4) VSV: VESfEISrIV (semivowel formed during braking by the external intercostals, released during action by the internal intercostals: /gowɔhn, bɪyə, čEYə/).

37.472. Thus, for all non-obstruent phonemes with upper-channel activity near the end of one chest pulse and/or the beginning of the next, the 'S'-timings of §37.471 may be phonemically relevant in terms of transition distinctions.

37.473. Likewise, for languages with true hiatus phenomena, the 'S'-timings might be read as subsuming the 'V to V' articulatory movements—movements which are allophonic in terms of upper-channel criteria, but which might have phonemically distinct timings in relations to chest pulses.

SECTION 4

40. In Section 1 we surveyed the so-called supra-segmental phonemes of American English and noted the descriptions of their allophones. In Section 3 we examined the Stetson model of speech-processes, and noted particularly those processes which might be involved in the production of the supra-segmental allophones. In this Section the two topics will be brought together, point for point; and an attempt will be made to hypothecate, from the Stetson model, articulatory processes which could produce the observed features associated with each of the several supra-segmental phonemes. These hypotheses are only just hypotheses; later, in Section 5, their plausibility will be examined.

41. The supra-segmental phonemes to be examined are the following: the point-like pitch phonemes /⁴, ³, ², ¹/ (see §42); the point-like stress phonemes /', ^, ` , ˇ/ (see §43); the transition phonemes /#, ||, |, +/, with their contour-like pitch and stress features, including duration (see §44); the syllable, its nucleus and its boundaries, as criteria for the classification of phonemes as syllabics or non-syllabics, and as places of incidence of transition features (see §45).

42.1. THE POINT-LIKE PITCH PHONEMES. (See §12 above.) These are associated

with laryngeal activity, and Stetson's model provides for a pitch-control mechanism in the larynx. There are certain limitations on the occurrence of /⁴, ³, ², ¹/ and certain predictable allophones in association with various stresses. But /⁴, ³, ², ¹/ are autonomous phonemes; and the mechanism which produces them is essentially independent of the other speech mechanisms; see §31 above. Such overlappings of the permitted occurrences of pitch and stress allophones as are found in American English are primarily linguistic and only secondarily physiological. Thus, the pitch phonemes usually have higher allophones in co-occurrence with the greater stresses; and this implies a more vigorous activity of the laryngeal musculature during segments of more-than-minimal airflow through the larynx. This co-occurrence can be described as some kind of physiological balancing of obstacle and moving air; but that description is not an explanation. Some other language, which also had point-like pitch and stress phonemes, might display quite different associations between them, and quite different allophonic interrelations. For instance, such an other language might display lower pitch allophones with major stresses, or allophones of upward or downward shift during major stresses; the total speech mechanism could be trained to such habitual coordinations.

42.2. It is well known that there is no simple linear relation between voice-pitch and any one musculature. But for linguistic purposes it probably suffices to stop with the crude first approximation: higher pitch is associated with greater vigor of larynx-activity; the vocal bands are subjected to greater tension, the area of intermittent air-escape between them is reduced, and the frequency of their vibration is increased.

42.3. Further, the pitch-producing mechanism can function in many different forms of association with the other speech mechanisms, both sub-laryngeal and supra-laryngeal; the pitch mechanism is independent of the others, except for the limitations noted in §10.11 and §12.2 above: minimum air flow, minimum duration; voice.

43.0. THE POINT-LIKE STRESS PHONEMES. (See §13 above.) These are associated with some organ of air-expulsion—normally, pulmonic air. The sub-laryngeal mechanism is usually treated as a single organ; even so catholically inclusive a survey of vocal activity as Pike's *Phonetics* dismisses this mechanism with the description (89): 'The lungs acting as an initiator may press lung air outward.'

43.1. The Stetson model is much more complex. Indeed, the sub-laryngeal mechanism, for Stetson, comprises two of the four principal sets of speech-producing organs; see §31 above.

43.2. Stetson's abdomen-diaphragm model provides for two degrees of surge superimposed upon the overall controlled movement of the breath group: a 'foot-stress', apparently /[^]/, and a 'main breath-group stress', apparently //'. The model calls for as many feet as there are //s and /[^]/s in a breath group, and as many //s as there are breath groups in a 'phrase'.

43.3. What is the nature of these // and /[^]/ surges? As noted in §33.3 above, there are three possibilities: (1) extra intensity of abdomen contraction, (2) momentary relative relaxation of diaphragm resistance, or (3) both. I cannot read from *Mot. Phon.*² any clear indication of a choice among these three. A guess,

but only a guess, would be that extra abdominal action is present in both // and /[^]/, and that // also has diaphragm relaxation. Another guess would be that both // and /[^]/ have diaphragm relaxation, and // also has extra abdominal action.

43.40. Meanwhile, what of /' and /[˘]/ ? Most discussions of American English stress phonemes arrive at a formulation with four distinctive degrees of relative stress /', [^], [˘], /[˘]/; I know of no explicit analysis of any groupings within this set of four phonemes. Implicitly, though, there are hints that /[^]/ is a reduced //, or that // is a promoted /[^]/, or even that ['] and [[^]] might be regarded as co-allophones of a phoneme of greatest stress, their occurrence being complementary in terms of certain pitch patterns.

43.41. If those hints were pursued, we might label the four stresses: // maximum, /[^]/ major; /' minor, /[˘]/ minimum. The implied groupings would correspond to an observation that some beginning students in a phonetics course find it difficult to discriminate between // and /[^]/, and between /' and /[˘]/ (But be it noted that in English verse /' is just as available as /[^]/ to serve as a 'metrical accent').

43.42. The disposition of /[˘]/ seems fairly simple in terms of Stetson's model: it would correspond to a chest pulse, one stroke of the internal intercostals. (We will discover difficulties with this interpretation later; see §44.532 below. But let us accept it temporarily.) We noticed, §§37.10-473 above, that there are several modes of chest-pulse activity. Are all of these to be associated with /[˘]/ stress? If so, what is the physiological correlate of /' stress?

43.430. 'Minor stress' /' is not discussed in *Mot. Phon.*² The examples in which /' might be expected, *upspring* (89), *instate*, *institute* (111), *Zeep Ope will be pope* (120), have no action-potential lines for chest-pulse as distinct from abdomen-diaphragm action; the lines for 'air outside, air in mouth, trachea' are non-discriminatory. Hence we may and must consider the various possibilities for accounting for the distinctive minor degree of stress /'.

43.431. // might result from some kind of abdomen-diaphragm action, different from that of /' and /[^]/. What could this be? Either a lesser degree of /[^]/, or some different combination of the factors involved in // and /[^]/. These would be possible, but somewhat untidy, and also contrary to the naive classification of // and /[^]/ as 'accented', /' and /[˘]/ as 'unaccented'. Neither consideration is conclusive, but they suggest looking elsewhere.

43.432. A more attractive possibility would be to regard /' and /[˘]/ as different kinds of chest pulses: /' distinctively stronger, in some way, than /[˘]/. Stetson does not discuss this possibility: see §35.2 above. But it would appear to be possible that the distinction between /' and /[˘]/ could be produced by different modes of intercostal action: that /' could correspond to a controlled movement, /[˘]/ to a free ballistic movement: see §23 above. Inherently, a controlled chest pulse, with the external intercostals acting throughout, but so as to yield somewhat to the dominating action of the internal intercostals, would be a prolonged pulse; the continuing contraction of the internal intercostals would result in a 'stronger' speech segment—at least impressionistically because of

greater length, perhaps objectively as well. Query: is it possible to pronounce with artificial prolongation a /˘/ without replacing it with /˘/?

43.5. We set up as hypotheses, then, the following 'Stetson-model' scheme for the point-like stress phonemes: // maximum: an abdominal surge accompanied by momentary relative relaxation of the opposing diaphragm musculature; /˘/ major: either an abdominal surge or a momentary diaphragm relaxation; /˘/ minor: a controlled chest pulse, with dominance of internal intercostal over external intercostal action throughout the pulse; /˘/ minimum: a ballistic chest pulse, with the motion of the chest wall checked, braked, stopped toward the end by various mechanisms. (But see §§45.40-43 below.)

44.0. THE TRANSITION PHONEMES. (See §14 above.) The stigmatic features of these phonemes are of various kinds, all of a 'contour-like' rather than a 'point-like' nature.

44.01. Pitch; see §§12.3-4. Before ///, there is a rise from the last point-like pitch; before //, there is sustension at the last point-like pitch; before /#/ and /+/, there are no significant pitch stigmata. After breath intake there may be a characteristic quick rise. Across close transition, a pitch shift from one to another of the point-like pitches is effectively continuous; across any of the four 'junctures' a pitch shift is discontinuous.

44.02. Stress; see §13.5. Before ///, and //, there is prolongation at the last point-like stress, terminated by abrupt cutoff or a rapid decrescendo which is perceptually equivalent to abrupt cutoff. Before /#/ and /+/, there is a relatively gradual decrescendo, during which stress declines to silence. After all the transition phonemes (and after breath intake, which may follow /#/ and //), the first succeeding point-like stress sets in, at full intensity, with the beginning of utterance.

44.03. 'Juncture allophones of the segmental phonemes'; see §13.51 and §13.6. These are also contour-like in nature: involved are such allophonic variations as pre-pausal 'drawling, unreleased or unaspirated stops, partial unvoicing, lengthening', and post-pausal 'smoothly-beginning or glottal-stop-released vowels, short obstruents, aspirated voiceless stops, voiceless-to-voiced obstruents'. To a quantitatively lesser extent, similar allophonic variants are found before and after // and /+/. Obviously, all these descriptions of allophonic stigmata attributed to 'juncture-context' imply a comparison with the allophones which occur in close transition. Clearly, in close transition American vowels are not initiated by glottal stop, or drawled; voiced obstruents are voiced throughout; voiceless stops are released, though not conspicuously aspirated.

44.1. An attempt to correlate these complex allophonic stigmata with the Stetson model impels a search for ONE physiological constant which is basic to 'the totality of the phenomena' impressionistically associated with each of the 'junctures'.

44.2. There is probably a connection between these two sets of statements: (1) Before ///, and //, there is a pitch-contour component, rise and sustension respectively; before /#/ and /+/, there is no pitch-contour component. (2) Before ///, and //, there is prolongation of the last point-like stress, terminated

by abrupt cutoff; before /#/ and /+/, there is gradual decrescendo. — We might conclude that before // and /|/, some air-driving mechanism continues in operation; the supra-laryngeal mechanism is in effect frozen at the last articulation; the larynx is pitch-constant for /|/, frozen at the pitch of the last resonant, but pitch-rising for //.

44.3. There may be a connection between these two sets of statements: (1) Before /#/ and //, voiced obstruents are unvoiced fairly early, and stops are unaspirated and usually unreleased; before /+ and /|/, voiced obstruents are unvoiced late or not at all, and stops are unaspirated but often released. (2) Before /#/ and //, there is a pre-inhalation posture; before /+ and /|/, there is not. — We might conclude that for /#/ and //, some air-driving mechanism is 'turned off' as for the end of one operation; for /|/ and /+/, it is kept on a 'stand-by' status, driving air at a minimum rate, but still driving.

44.4. There is almost certainly a connection between these two sets of statements: (1) Alternative pronunciations of 'the same utterance' (i.e. one adjudged to have certain given morphemes in a given order, except for differences in pitch, stress, and transitions) usually involve differences in both pitch and /#, ||, |; alternative pronunciations of 'the same part of the same utterance' usually involve differences in both stress and /+ vs. close transition; in the Trager-Smith terminology, there are clause intonation patterns, and word and phrase superfixes. (2) In the Stetson terminology, there are breath-group and foot sequences.

44.50. Some hypotheses to relate the transition phonemes to articulatory activity would be the following.

44.51. The pre-intake transitions, /#, ||, are produced by characteristic actions of the abdomen-diaphragm musculature, in the timing whereby the diaphragm's braking assumes dominance over the abdomen's driving. With /#/ , the abdomen driving stops and the diaphragm braking absorbs the momentum; with //, the abdomen driving continues somewhat longer and the diaphragm braking is more sudden.

44.52. With /|/, there is reduction or momentary cessation of abdomen driving, but no dominance of diaphragm braking. /|/ is not a pre-intake transition and hence is not followed by positive inhalation action by the diaphragm.

44.531. With /+/, the controlling mechanism is the intercostal musculature. The crucial feature here is one of timing as between the intercostal and the supra-laryngeal mechanisms; see §§37.40–471 above. The decisive feature is the relation between two shifts: (1) from external intercostal dominance (or air-pressure braking by a C-obstruction) to internal intercostal dominance; (2) the removal of some C-articulation in the supra-laryngeal tract. If that C-removal takes place early in an internal intercostal pulse, there is close transition. If C is removed before internal intercostal action, there is C+. If C-removal comes after the internal intercostal pulse is well advanced, there is +C. If C-formation is not accompanied by external intercostal braking and if C-removal comes late in the internal intercostal pulse, there is C + C.

44.532. A special problem connected with /+ is exemplified in §14.221 above: 's+cól+today #, 'f+'ts+sqüärelly | 'n+frónt #. Reasoning from the

hypotheses of §37.4 and §44.531, we should expect to find /+/ whenever a C-removal occurs after an internal intercostal pulse is well advanced. If we find /s+k/ in 's cool today but /sk/ in school today, we should expect that /s+k/ corresponds to a sequence: /s/-formation, ii pulse completed, /s/-removal, /k/-formation, ii pulse, /k/-removal; and we should expect that /sk/ corresponds to a sequence: /s/-formation, ii pulse initiated, /s/-removal, /k/-formation, /k/-removal, ii pulse completed. In other words, /s+kuwl/ would have two internal intercostal pulses, /skuwl/ would have one. With /s+kuwl/ the second internal intercostal pulse would override the first, the /s/ articulation being removed and the /k/ formed before, and /k/ removed during, the second internal intercostal pulse. With /skuwl/ there is only one internal intercostal pulse, with the /s/-removal, /k/-formation and /k/-removal all occurring during that pulse: /s/ releasing much of the built-up pressure, /k/ releasing some residue, and /uw/ driven by the surplus contributed by an abdomen-diaphragm surge corresponding to '//. (In general, we should expect to find maximum aspiration of /p, t, k/ when they occur immediately after /#, ||, |, +/ and with a major or maximum abdomen-diaphragm surge.)

44.533. The hypotheses of §§44.531-2 can be adapted to /f+ts+skwéhrlɪy | n+frént #/.

44.534. The internal intercostal mechanism postulated for /s+kuwl/ is not far-fetched in terms of the Stetson model, however disturbing it may be in terms of conventional operations with undefined 'syllables'. If we keep strictly to physiological articulatory mechanisms, there is no difficulty in operating with an internal intercostal pulse released and arrested by /s/, followed by a second internal intercostal pulse released by /k/. Likewise, with /skuwl/, there is nothing mysterious about a single internal intercostal pulse during which /s/ is removed and /k/ is rapidly formed and removed. With /s+kuwl/'s two pulses we expect /s/ to be longer (as release AND arrest) and /k/ to be aspirated (as sole C-release). With /skuwl/'s one pulse we expect /s/ to be shorter and /k/ to be very short and unaspirated. And so they are.

44.535. The chest-pulse mechanism required for /+/ within a consonant sequence without a vowel is described in §22.43 and §37.45 above: what Stetson calls 'the adventitious syllable' and what we have postulated as a sequence marked by phonemic /+/ transition.

44.6. The mechanisms postulated for /#/ and /+/ have certain similarities, which is not surprising, since both of these transitions are characteristically gradual-decrescendo and 'pitchless'. With /#/ we assumed (§44.51) that the abdomen driving stops and the diaphragm braking takes up the momentum; any resumption of speech requires a fresh abdomen driving, with supra-laryngeal articulations adjusted as air-flow beings. With /+/ we assumed (§44.531) that the internal intercostal driving stops and the momentum is taken up by either the external intercostals or an arresting C-obstruction; the resumption of speech requires a fresh internal intercostal driving, with supra-laryngeal articulations adjusted as air-flow begins.

44.7. We set up as hypotheses, then, the following 'Stetson-model' scheme for the transition phonemes: /#/ terminal fade: abdomen action stops and di-

aphragm braking gradually takes up the momentum; /||/ terminal rise: abdomen action continues, then stops, and diaphragm braking (simultaneously and forcefully) takes up the momentum, laryngeal pitch rises; /|/ medial sustain: abdomen action reduced or stopped, without effective positive diaphragm counteraction; /+/ open transition: removal of a C-obstruction before or after, but not during the build-up phase of an internal intercostal pulse.

45.0. THE SYLLABLE. (See §15 above.) Contemporary descriptions of American English phonemics operate in various degrees with a 'syllable' which has a nucleus, which has boundaries, which has phases so that the same supra-laryngeal articulation is sometimes syllabic and sometimes non-syllabic, at the beginning or end of which there are phenomena characterizing transition phonemes. Let us consider these 'properties of the syllable' in terms of Stetson's model.

45.10. The syllabic nucleus. This is a segment which neither releases nor arrests a chest pulse. If the first audible portion of a chest-pulse is a nucleus, the pulse is entirely chest-released, and sound-production is wholly dependent on the build-up of air pressure by internal intercostal action. If the last portion of a chest-pulse segment is a nucleus, the pulse is entirely chest-arrested, and the external intercostals must act to take up the momentum.

45.11. This interpretation of 'syllable nucleus' would agree with Trager-Smith and with Stetson. For Trager and Smith (see §15.0 above), 'non-syllabicity' appears to be the primitive term, and a syllabic is a segment which is not non-syllabic. For Stetson (*Mot. Phon.*² 37a), 'the consonant has a function in the movement of the syllable; it constricts the vocal canal to delimit the chest pulse; the vowel on the other hand is a shaping movement to emit the chest pulse.'

45.12. But this interpretation suffers from the common weakness of negative definition: somewhere in the chain of equations the devil of 'necessary but not sufficient' creeps in.

45.131. What of /s+kuwl/; is /s/ a syllabic nucleus? No, according to this interpretation; for /s/ releases and arrests a chest pulse.

45.132. What of /bɪɹ/ if, after the /b/ segment, there is a uniform supra-laryngeal articulatory posture; is /ɪɹ/ a syllabic nucleus? According to this interpretation, /ɪɹ/ would be a nucleus, for the pulse is entirely chest-arrested. So also in /bɪɹt/; for there /t/, not /ɪɹ/, arrests with or without external intercostal action.

45.133. But what of /yes/; is /ye/ a syllabic nucleus? By this interpretation, /ye/ would be a nucleus. So would /yey/ in /yeyl/; and so would /ney/ and /rey/ in /neyl/ and /reyl/. (Possibly relevant query: Could any of these forms be released with a glottal stop?)

45.134. And what of /ɹən/, /rowl/, /mɪɹn/; are they nuclei?

45.14. The interpretation of 'syllabic nucleus' as 'a segment which neither releases nor arrests a chest pulse' would agree with analysis so far as the examples in §§45.131-2 are concerned, and would conflict so far as the examples in §§45.133-4 are concerned.

45.2. The syllable boundaries. Stetson's model defines the passage from one syllable to the next in terms of a change from braking dominance (by external intercostals and/or C-barrier) to driving dominance by the internal intercostals,

either by a free stroke continued by chest-wall momentum or by a controlled movement. The various coordinations of this shift from braking to driving dominance with the supra-laryngeal articulatory shifts can be phonemically distinctive, in terms of close vs. open transition.

45.30. Phases of the syllable: syllabic vs. non-syllabic. The distinction between syllabic and non-syllabic status is an essential one in the derivation of such a system of English phonemes as Trager and Smith present. If the 'second elements' of *eye* and *A* are to be co-allophones of a phoneme /y/, although the second element of *eye* is phonetically similar to /ɛ/ in *let* and the second element of *A* is phonetically similar to /ɪ/ in *lit*, it must be on grounds that these second elements are somehow a different order of segments from the first elements or the elements in *let* and *lit*. That difference in order is called 'non-syllabicity'.

45.31. We sketched above, in §45.10, a possible criterion for distinguishing between non-syllabics and syllabics. Another possibility—a procedure implied by Hjelmslev in his paper *The syllable as a structural unit* (*Proc. 3d Int. Congr. of Phonetic Sciences* 266–72)—is to accord a certain primacy in analysis to the sub-laryngeal activities. (Such a primacy may seem unconventional; and it is, in terms of the history of linguistic analysis. But in practical speaking, the sub-laryngeally determined phonemes resist mishearing more successfully than the 'segmental' phonemes, including 'vowels'.) Thus we would first establish, for American English, the repertory of the significantly different patterns of pitch, stress, and transitions. The sites of the point-like pitch and stress phonemes would be found to coincide with occurrences of some but not all the supra-laryngeally determined phonemes, or with some but not all of the occurrences of their allophones—allophones from the supra-laryngeal point of view.

45.32. Thus, occurrences of a phoneme /b/ are not found as the site of a point-like stress or pitch phoneme. Mid-front vowels, like those in *let* and the beginning of *A* and the end of *eye*, are the sites of point-like pitch and stress phonemes in *let* and *A* but not in *eye*. Hence the middle of *let* and the beginning of *A* 'bear' or 'are borne by' point-like pitch and stress phonemes. The beginning of *let* and the end of *eye* do not 'bear' or are not 'borne by' such point-like phonemes.

45.33. Such a formula could define a 'syllabic', though it might be asked whether that would be the most appropriate term for it. If it is considered appropriate, one might then derive some corollary definitions of 'syllable', or 'full syllable', or the like. That formula would assert that there are no 'syllabics' in /f+ts+skw/ in the stretch /f+ts+skwéhrlɪy/; for there are only two sites of point-like pitch and stress phonemes. Perhaps such an assertion is just what we want, for certain purposes of analysis.

45.34. Another consequence of the procedure sketched in §45.31 would be a positive significance of /˘/ 'minimum stress'; for the difference between minimum stress and no stress would be a crucial criterion for the determination of some phonemes: the site of minimum stress is a 'syllabic'; the site of no stress is a 'non-syllabic'. To be sure, 'no stress' would have to be defined phonemically as 'no degree of stress variable in the context'.

45.35. Thus, the treatment of 'syllabic r, l, m, n, ŋ' would start from the ob-

servation that *butter*, *bottle*, *open* [-pm], *button*, *pumpkin* [pəŋkɪŋ] or [pəŋʔŋ] have the stress form /' ˘/. The phonemic interpretation as 'syllabic r, l, m, n, ŋ', or as /ə/ or /ɪ/ with following /r, l, m, n, ŋ/, by arguments from free variation, would follow the familiar procedures.

45.40. The suggestion of §45.31 that /˘/ be regarded as a positive degree of stress would make it possible to assert that the sequences /s+k/ in 's *cool today* and /f+ts+skw/ in 'f 'ts *squarely in front* lack 'syllabics' since they lack a stress phoneme. But it would also involve the necessity of reconciling some discrepancies as between two other hypotheses: (1) in §43.42 and §43.5, it was hypothesized that the physiological correlate of /˘/ is 'a ballistic chest pulse'; (2) in §§44.531-533 and §44.7, it was hypothesized that /+/ presupposes two chest pulses. But in /s+k/, /f+ts+skw/, and the like, one or more of the chest pulses does not correspond to /˘/.

45.41. Stetson evades this dilemma by calling the stressless pulse segment 'a preliminary silent syllable' or 'an adventitious syllable' (see §22.43 above). Further, the drawings (*Mot. Phon.*² figures 72, 74, 75, 76) offer no direct evidence on either intercostal or abdomen-diaphragm action.

45.420. There seem to be two possible solutions; in the absence of instrumental evidence, there is at present no basis for preferring one or the other.

45.421. The stressless chest pulse occurs when some one obstruent functions as BOTH RELEASE AND ARREST of a chest pulse. Thus, in /s+kuwl/, the first pulse is released and arrested by /s/, the second pulse is released by /k/. In /f+ts+skwehrlɪy/, the first pulse is released and arrested by /f/, the second pulse is released by /s/ (as the second phase of a compound—not abutting!—consonant group) and also arrested by the same /s/. In /v+menɪy/, /v/ releases and arrests the first pulse. This interpretation would set up, as a requirement for an occurrence of a stress phoneme, some interval between the releasing and arresting Cs of a chest pulse. It is thus essentially a corollary of §45.10, where a 'syllabic nucleus' was defined negatively as a segment not releasing or arresting a chest pulse, plus the definition of a nucleus as a segment 'bearing' or 'borne by' a stress phoneme (§45.31).

45.422. Another solution: the stressless chest pulse occurs when there is no upward displacement of the diaphragm, when the sole driving force is the internal intercostal pulse. This would require assuming some immobility of the diaphragm—either by resistance of the abdominal viscera or by some compensatory abdominal fixation—to account for the expulsion of air via the upper vocal tract. Or perhaps such stressless chest pulses occur during a period of minimum upward displacement of the diaphragm: a slow or delayed beginning of a breath-group action by abdomen-diaphragm, or a premature ending (as in languages with various 'reductions' in pre-pause segments), or—when medial in a breath group—a temporary minimum of the net controlled movement (see §21.4, §22.1, §23.1).

45.43. Neither interpretation is neat. The untidiness of the first (§45.421) is that it involves a supra-laryngeal factor in the stress system. The second (§45.422) complicates the abdomen-diaphragm model by introducing a fourth mode of

activity in addition to those described in §33.5—and that to account for a relatively rare free-variant pronunciation. Until some experimental data are in, the problem of /+ / preceded and/or followed by stressless chest pulse must be listed as one of the loose ends of articulatory phonetics.

SECTION 5

50. Section 4 was a blueprint of possible bridges between the phonemes of stress and transition and the Stetson model. It is quite certain that it requires corrections: Stetson and the analysts who derived the phonemes were not working within the same frame of assumptions nor with the same goals. To what extent is that blueprint sheer fantasy?

51. The question is essentially: How reliable are Stetson's measurements, how sound are his conclusions, and how pertinent are both to a phonemic analysis of stress and transition in American English? For, whatever lacks we may find in Stetson's expositions, it is still Stetson and not the analysts who provides discussible statements. The analysts have almost entirely ignored the mechanisms which produce the phonemic distinctions of stress and transition, and—since those distinctions are predominantly relative rather than absolute—the analysts have not been able to gather the large body of acoustic data which would be necessary to quantify the psycho-social judgments. Those judgments are crucial, of course, in phonemics; and many analysts have been content to stop with them as a description of speech behavior.

52.0. How valid, then, is the Stetson model of speech, quite apart from the attempts I have made here to connect it with the 'supra-segmental' phonemes?

52.1. Were Stetson's measurements sufficiently accurate? I cannot swear to his tracings; but for linguistic purposes the absolute scale of time and magnitude factors is much more than sufficiently small, so that a wide margin of error could be tolerated.

52.2. How valid are Stetson's interpretations of his measurements? Here the situation is less cheering. Much of Stetson's research was based on data from tambours and thistle tubes recording through the body walls, or from a gastric balloon, or from laryngectomized or tracheotomized subjects. Relatively few oscillograph records of action potentials during speech motions appear to have been used. The disadvantages of mechanically made records are obvious: the speaker is seriously constrained by the sheer bulk of apparatus and by an unnatural total posture; the loss of detail and the blurring of time-factors may be considerable, in view of the speed and small absolute magnitude of some of the movements involved, when the records are made through cushions of flesh and by the activation of instruments whose inertia is considerable in relation to the forces under study. Heffner's characterization (*General phonetics* 74), 'rather fussy laboratory techniques', is not an overstatement. It is clearly desirable that we have many more action-potential records of sub-laryngeal speech movements. — However, the existing records seem adequate to establish the minimum fundamental data needed for an articulatory model of the supra-segmentals: the

occurrence of intercostal pulses, variously synchronized with supra-laryngeal movements; the occurrence of abdomen-surges and (probably) surges of distinctively different intensities.

52.30. How pertinent are Stetson's measurements and his conclusions for our purposes in this discussion? The distressing fact is that, in all his laborious recording and measuring work of several decades, Stetson did not set up one crucial experiment which would establish or disprove the hypotheses advanced in Section 4 of this paper.

52.31. Stetson was concerned with a unit he called 'the syllable'. He was only incidentally interested in stress and transition phenomena. He operated with at most three levels of stress, apparently // breath-group stress, /[^]/ foot stress, /[˘]/ no stress—ignoring /[˙]/. He studied only two kinds of transition, the contrast between compound and abutting consonants, or, in his studies of 'changes with varying rates', the contrast between arresting and releasing functions of a C.

52.32. Most disastrous for linguistic purposes is the kind of material used by Stetson: single words, or nonsense syllables, or nonsense-like phrases, repeated at increasing or decreasing rates. Very rarely, a continuum of a sentence-like kind was used, and some of the phenomena of normal American English speech could occur. Thus it is recorded (*Mot. Phon.*² 120) that 'A series like *Zeep Ope will be pope* is likely to be given as two distinct breath groups, *Zeep' Ope' / will' be pope.*' Page 121: 'The prescribed formula *pop up' a pop' up* was uttered ... *pop' / up' a / pop' up* ... in 18 of the 19 cases. In 10 of the 19 cases, the initial one-syllable foot *pop* is marked off by an arresting consonant, the adventitious double *pop-pup* is avoided to keep the verbal *up* quite clear.' — Also, Stetson gives useful discussions of pairs differing by a transition factor, such as *top pick* compared with *topic* (but without mentioning the difference in stress!). But most of the figures represent such utterances as *peb' pep, peb pep'* or *te te te' te, te te' te te* or *runnin' 'n' neighin'* or *Z is Z or Lil' 'll lie low*.

52.33. As the quotations in the preceding paragraph show, Stetson uses an apostrophe to mark 'accented'; no distinction is made between various degrees of stress. Presumably this marking was a cue for the speaking subject. There is no indication that simultaneous acoustic records exist; so we will never know just what the speaking subject actually said, so far as stress and transitions are concerned. Most of the measurements far antedated the analysis of stress and transition phonemes; so Stetson could not, even if he had wanted to, present a full transcription of the actual speech.

52.34. A serious deficiency of the drawings, for our purposes, is the general omission of data dealing with abdomen-diaphragm action and its synchronizations with intercostal action. This is understandable in view of the intricate machinery required, but some important questions about the production of stress phonemes and terminal phenomena can be answered only by experiments with distinct recordings of the several major sets of articulatory organs. The lines in Stetson's drawings showing air pressure in the mouth or (in the operated subjects) in the trachea represent the indiscriminated total result of abdomen-diaphragm AND intercostal action. Further, such measurements leave unanswered a question of fundamental theoretical importance: How much of the pressure

variation is the result of contraction of the thoracic cavity, how much the result of obstruction in the supra-laryngeal tract?

53.0. What plausibility, then, resides in the hypotheses put forward in Section 4 of this paper?

53.1. Obviously, very little, in specific detail. The lack of crucial experiments, of realistic linguistic material, of adequate transcription of utterances—these lacks forbid any claim that a particular supra-segmental phoneme has been correlated with a particular muscular behavior.

53.2. In general, though, there must be some validity in the overall model. For all its inadequacies for linguistic purposes, Stetson's material points toward two sub-laryngeal mechanisms, each capable of distinctively different modes of action, and both capable of various time-sequence relations with each other and with the larynx and the supra-laryngeal mechanisms. If phonemes of stress and transition exist, they correspond to articulatory habits; they are classes of vocal behavior by human speakers; that behavior requires muscular activity; and the abdomen-diaphragm and the intercostal musculatures are known to be active during speech.

53.3. Thus the hypotheses of Section 4 are more a program than a report. They represent one person's guesses as to the most promising possibilities to be tested first, in a project to specify the articulatory model of the supra-segmental phonemes. Quite certainly the objective evidence, when it exists, will show that some of the guesses were quite wrong, others over-simple or over-complex. Perhaps it will be found that diaphragm relaxation is the basic correlate of the breath group and its terminal transitions and that positive abdominal action is the surge mechanism, or that the rectus abdominis and the lateral muscles function independently; perhaps the triangularis sterni as well as the intercostals functions distinctively in chest pulses. At all events, it seems clear that more detailed study of the sub-laryngeal mechanisms, using material which is linguistically analyzed and notated, is desirable to bring the supra-segmental phonemes out of the realm of subjective impressionistic identification, of dogmatic description, and of often unprofitable controversy.

LARYNGEALS AND THE INDO-EUROPEAN DESIDERATIVE

JAAN PUHVEL

Harvard University

This is an attempt to suggest briefly a systematic approach to a number of desiderative or future formations in various IE languages. Taking laryngeals into consideration, I hope to show their possible inclusion under a single formula. No exhaustive study of all variants or subtypes is intended, and no attempt is made to trace the working of analogy which was certainly involved in formations of this kind. The main object is, however, to disqualify some of the more awkward appeals to analogy which have seemed necessary so far.

The formations included are the Sanskrit desiderative and future, various types of Greek future, and the Old Irish strong futures. To the same general category belong further desiderative presents like Skt. *rdkṣati*; Gk. *αἴξω*, *ἀλέξω*; Lat. *uīdō* < **ueidsō*, *quaesō* (cf. also Lat. *faxō*); Goth. *fraliusan*, etc.; further, the Lithuanian future (*būsiu*), Old Slavonic relics, and possibly Oscan-Umbrian formations.¹ The ultimate affinity of the *s*-aorists is not impossible.

The Sanskrit desiderative² normally contains (1) accented reduplication with weak-grade vowel (*i*, *u*), (2) weak grade of the root, (3) *s*, and (4) thematic flexion, e.g. *vi-vrt-sati*, root *vrt-*. In roots with final resonant, however, as has been observed by Schulze (*Kl. Schriften* 101 ff.), the picture is more complicated: e.g. *ci-kīr-ṣati* (*kīr-*), *jī-ghān-sati* (*han-*). *īr* and *ā(n)* are the normal Sanskrit results of *ṛ*, *ṝ*, which led to the explanation that 'originally *s* was separated from the final consonant of the root by *a*, which combined with the preceding liquid or nasal to give de Saussure's *ṝ* ...'.

In modern terminology this means that *r*, *ṛ*, and the other syllabic resonants (cf. *jigīṣati*, *ṣuṣrūṣati*) were followed by a laryngeal, the loss of which brought about their lengthening to *ṝ*, *ṝ̄*. Accordingly the desiderative element was not *-s-* but *-Ḥs-*. This odd-looking formative may be the weak grade of IE **Ḥes-* (with *e*-colored laryngeal) 'to be', and the whole ending *-Ḥseti* may be a thematic inflexion of the same root with the accent originally on the thematic vowel, i.e. **Ḥēs-ti*, **Ḥs-ētī*, which is normal. We get thus *cikīrṣati* < **q̄i-q̄r-Ḥseti* 'he is (for) doing' or 'he is doer'; *jīghāṇsati* < **gh̄i-gh̄n-Ḥseti* 'he is (for) slaying' or 'he is slayer'. This would be paralleled closely by the periphrastic future in Sanskrit: *kartāsmi* 'I am doer'. The root expressed the verbal notion with great flexibility, probably originally as an action noun, but also, perhaps through a primitive synthetic bahuvrīhi development,³ as agent noun (Skt. *dviṣ* 'hostility' or 'enemy'). Cf. further the Sanskrit periphrastic perfect *corayāmāsa* beside *corayāmcakāra*. Another later semantic parallel is provided by the Greek desiderative; a time-honored approach sees in *δράσειων* (Soph. Aj. 326) a primitive or analogical **δράσει ὶων* 'going for doing', enlarged into full conjugation and

¹ For these, cf. Magnien, *Emplois et origines du futur grec* 285 ff. (Paris, 1912).

² Whitney, *Sanskrit grammar* 372 ff.

³ Cf. Puhvel, Indo-European negative composition, *Lg.* 29.14 ff. (1953).

expanded as a type. Many of the cases found, especially in the older authors, show participial forms which were the probable starting point: *ὄψιόντες* (Il. 14.37), *ἐργασίων* (Soph.), *γελασείοντα* (Plato), *ἀπαλλαξείοντες*, *ναυμαχῆσειοντες*, *συμβασειόντα* (Thuc.).

The problem is similar to that of the Lat. *-b-* future, for which the most satisfactory explanation still is a periphrastic formation with IE **bheṷ-H-* 'be, become', i.e. *vidē-bit* from **bh(u)ṷet(i)*, subjunctive of the non-thematic root aorist seen in Skt. *bhūvat* (indicative *ābhūt*), and 3rd plural *-bunt* < **bh(u)ṷont(i)*, Skt. *bhūvan*, OCS *bq*. It is therefore conceivable that **-Hseti* might likewise be a primitive short vowel subjunctive.

The connexion of the Old Irish strong futures with this Sanskrit desiderative has been pointed out by Thurneysen.⁴ The development may now be reconstructed as follows: **gh^high^hηHseti* > PKelt. **gignāseti* > OIr. *génaid* (fut. of *gonaid* 'wounds'); **gigrHseti* (Skt. *grīṇāti*, Gk. *γῆρς*) > PKelt. **gigrāseti* > OIr. *géraid* (fut. of *gairid* 'calls').

This approach may be clarified further when applied to the future formations in Greek.⁵ It is commonly held that the Greek future is mainly descended from the IE desiderative.⁶ The desiderative was not necessarily reduplicated and could have normal grade of the root, as is shown by Vedic forms like *gṛṣamāṇah* 'wishing to hear'. Greek futures are not normally reduplicated, but forms like Hom. *δεδέξομαι*, Att. *κεκλαύσεταί* show the potential presence of reduplication.

Positing *-Hseti*, plur. *-Hsonti* as the desiderative ending, we distinguish the following varieties:

1. Verbal roots in a final nonsyllabic generally lost the following laryngeal between two nonsyllabics: **Heḡ-Hsonti* > *ἄξουσι*; **Hes-Hsetai* > *ἔσσεται*.

2. Roots in final *r, l, m, n*: **Hmen-Hsonti* > **μενεσσοντι* > *μενέουσι* > *μενούσι*; **ger-Hsonti* > **κερεσσοντι* > *κερέουσι* > *κερούσι*, where (μεν)έ(ουσι) is the reflex of a schwa vowel which shows a tendency to haunt especially the interspace of laryngeal and resonant, like the 'prothetic vowel'. Its presence may not be constant, for Hom. *φθέρσω* and *κέλσω* beside Ion. *φθερέει*, if ancient, point to an alternative situation where the laryngeal was simply lost between nonsyllabics.

Thus this type is not 'generalized from dissyllabic stems', but has a primary origin in early IE phonology. The result of a 'dissyllabic stem' is here in no wise different, but there was original collision of laryngeals: **kor.H-Hsonti* > **κορεσσοντι* > *κορέουσι* > **κοροῦσι* (reformation *κορέσουσι*). Likewise in 'dissyllabic stems' with some other nonsyllabic as root-final: **pet.H-Hsontai* > **πετεσσονται* > *πετόνται* > *πεσύνται*.

This may have certain parallels in the Sanskrit desiderative, in the oldest appearances of the 'union-vowel *i*'. A Vedic form like *jigamīṣati* (Tāittirīya Saṁhitā, also *ājigamīṣanto* in Aṣvalāyanasya Gṛhyasūtra 4.1.3), beside *jigāṁṣati*, is probably **g^hig^hηm.Hseti* vs. **g^hig^hηHseti*, like Gk. *θάνατος* : *θνητός* <

⁴ Thurneysen, *A grammar of Old Irish* 414 (Dublin, 1946).

⁵ For a general picture, see Magnien, op.cit. (fn. 1) and *Les formes du futur grec* (Paris, 1912); Chantraine, *Morphologie historique du grec* 289 ff. (Paris, 1945); Buck, *Comp. grammar of Greek and Latin* 228 ff. (Chicago, 1948).

⁶ Cf. Meillet-Vendryes, *Grammaire comparée des langues classiques* 192, 210 ff. (Paris, 1948).

**dhun̥Htos* : **dhun̥Htos*.⁷ Certain 'dissyllabic stems' also have early Vedic forms of similar nature: *jīṣvīṣati* (Vājasaneyi Samhitā) < **gʷigʷiH̥u.H̥seti*, beside *jūjyūṣati* (Catapatha Brāhmaṇa) < **gʷugʷiH̥u.H̥seti* from IE **gʷéj̥-H̥u-*, **gʷi-H̥-éj̥-*, **gʷi-H̥-u-*; *pipatīṣati* (Atharva-Veda) < **pipet.H̥H̥seti*, cf. **πτεσσονται* above; also *pipatsati*, either with the laryngeal lost between nonsyllabics as in *prattas*, or by analogy from the 'hyper-regular' present *pātati*. Other early cases of *i* are seen in the ÇB (cf. Whitney 1031b). Some, like *cikramīṣati*, *jijaniṣati*, *jigrahīṣati*, are from 'dissyllabic stems', and may thus reflect historical continuity of phonematic development. Others are probably already victims of that analogical parvenu of Sanskrit grammar, the generalized 'union-vowel *i*', which has its origin in similar cases.

This is not the place to consider fully the phonematic problems concerning the relationships of laryngeals, resonants, and schwa. The matter is fundamental to IE phonology, and I hope to return to it later. We may note here, however, that an opposition like *-nm̥H̥-* : *-n̥H̥-* is clearly a genuine ablaut relationship, and that the phoneme *̥*, as the dominating syllabic element, determines the allophonic conditions of both resonant and laryngeal. This recognition seems to contradict the assumption of a 'vocalized laryngeal': in *ἀνδρα* < **H̥nr̥n̥*, the **H̥n̥* can be only this succession of three phonemes; otherwise it ought to have been **H̥n̥* 'prior to vocalization', but then there would be no vocalization before a syllabic. If the laryngeal were somehow 'syllabic from the beginning', then laryngeals ought to be altogether more vowel-like than resonants, which is unlikely. (According to IE allophonic laws resonant before nonsyllabic is nonsyllabic only when preceded by a phoneme of greater syllabic force, i.e. a vowel; *ent* and *int* are possible, but not *int̥*.)

3. A final syllabic of the stem was lengthened by *H̥*; e.g. *ἀνέξουσιν*, future of *ἀνέω*, which is probably double desiderative: **H̥euḡ-H̥se-H̥sonti*; causatives like *σοβήσουσι* (*σοβέω*, Skt. *tyajayati*) < **t̥iogʷe-H̥sonti*. In denominative verbs like *φιλέω*, *τιμάω*, *δηλόω* the present is formed by the denominative element *-i-* from nominal *φίλος*. The future *φιλήσουσι* is unconnected with this formation and becomes fully intelligible only when analyzed **φιλε-H̥sonti* > *φιλή-σουσι* 'they are (for) loving' or 'they are lovers'.

4. The 'Doric' future is probably not a mere conflation of *-σω* and *-εω*. It is perhaps to be connected with the Lithuanian future in *-siu* and the Sanskrit future in *-sy-*; all may be *-i-* extensions of the non-reduplicated variety of desideratives, Greek in the form **-(e)H̥se-i̯onti* (*-σεοντι*): Hom. *ἔσσειται* < **H̥es-H̥seietai*; Sanskrit from **-(e)H̥s-i̯onti* (e.g. *kariṣyānti*).

⁷ Those who prefer something like *-m̥H̥-* to *-nm̥H̥-* are working with theoretical entities of Hirt and Cuny-Couvreux instead of Brugmann-Edgerton and Kurylowicz-Sturtevant. Certain features of my reconstructions are somewhat controversial, such as the allophone *-n̥n̥-* and the long resonant *-ñ̥-*. Their use represents no belief on my part in their particular phonetic probability or even phonematic superiority, but they are employed as the theoretical symbols that I find most useful in systematizing the IE phonematic relationships.

FURTHER OLD NORSE SECONDARY FORMATIONS

ALBERT MOREY STURTEVANT

University of Kansas

Certain Old Norse secondary formations were discussed in *Lg.* 29.26-33 (1953). Here I shall continue the discussion by adding examples which at the time of my previous article I had not yet investigated.

1. THE LENGTHENING OF THE VOWEL *e* IN OLD ICELANDIC *eta* > *éta* 'TO EAT'. Initial short vowels sporadically suffered lengthening in Old West Norse, more often in Old Norwegian than in Old Icelandic.¹ This lengthening represents the beginning of a trend which later extended to a large number of words. In the few examples recorded in Old Norwegian and Old Icelandic there is no plausible explanation of the lengthening except as due to the initial position of the vowel with accentual variation, a purely phonetic process. Old Norwegian *af* > *df* 'of', *at* > *dt* 'at', *akr* > *dkr* 'field' are on a level with Old Icelandic *ek* > *ék* 'ego', *eta* > *éta* 'to eat'. Noreen, nevertheless, assumes that the lengthening of the initial vowel *e* > *é* in *eta* > *éta* does not represent a phonetic process but was probably of analogic origin, the long vowel *é* following the pattern of the long vowel *á* in the anomalous preterit singular form *dt*, a type of borrowed lengthening.² This assumption is unfounded, primarily because Noreen has disregarded the lengthening of the initial vowel *e* > *é* in words which do not belong to a verbal ablaut system, such as the pronoun *ek* > *ék*, which cannot be of analogic origin. There is therefore no reason to assume that the forms *ék* and *éta* are on different levels. Furthermore, Noreen has not justified his assumed type of borrowed lengthening. He has not answered the question WHY this borrowing should have occurred, but has evidently based his assumption upon the fact that the vowels in *éta* and *dt* are both long. But this fact is no evidence that one vowel (*é*) became long because the other (*á*) was long. If Noreen's assumption were correct, we might expect the non-initial vowel *e* in the type *vega* 'to slay', *fregna* 'to ask' (of the same ablaut series) to have likewise suffered lengthening in imitation of the long vowel in the preterit singular *vá* and *frá*. That this never occurred is evidence that the vowel in *eta* was lengthened because of its initial position, parallel to *ek* > *ék*.

2. *Ftandr*, NOMINATIVE-ACCUSATIVE PLURAL OF *ftandi* 'ENEMY'. The substantive *ftandi* represents the present participle of the verb *fta* 'to hate', substantivized as a consonantal *nd*-stem. The nom.-acc. plural has the ending **-iR*. We should therefore have expected, instead of *ftandr*, a form **ftendr* (< **ftandiR*) with *i*-umlaut of the vowel **a* in the suffix. It is rightly assumed that the vowel *a* in *ft-a-ndr* was borrowed from the singular paradigm *ft-a-nd-i*, *-a*.³ But this is the only recorded example of an Old Icelandic *nd*-stem with such borrowing.

¹ Noreen, *Altisländische und altnorwegische Grammatik** §126.2.

² Noreen §498 Anm. 1: 'Neben *eta* tritt *éta* (*iéta* §103), wo die vokallänge wol aus dem prät. *dt* (got. *fr-ēt*, lat. *ēdi*, gr. *ἔδ-ῆα*) herübergenommen ist, schon seit dem anfang des 13. jahrs. auf'.

³ The leveling of the vowel in the plural *ftandr* in favor of the *a* in the singular *ftand-i*,

First of all, it should be noted that *fīandi* is the only *nd*-stem in which an *i* of the stem syllable is followed by the *a* of the suffix. This combination *ia* was later shifted to *id* (*fīandi* > *fjāndi*), just as in the infinitive *fīa* > *fjā*. The original nom.-acc. plural **fi-endr* could not follow this pattern, since the front vowel *i* was subject to this type of shift only before the back vowel *a*, not before the front vowel *e*. By restoring the back vowel, the nom.-acc. plural *fī-a-ndr* (> *fjāndr*) was brought into conformity not only with the singular paradigm *fī-andi* (> *fjāndi*) but also with the derivative verbal forms *fī-a* > *fjā*, *fjāða*, *fjāðr*.

In support of this hypothesis is the fact that the vowel *a* of the nom.-acc. form of the *nd*-stems did not occur in Old Icelandic except in *fī-a-ndr*, where the combination *i* + *a* occurred, though traces of this leveling are recorded for Old Norwegian.⁴ The contracted nom.-acc. *fjāndr* was on a level with the nom.-acc. form of the monosyllabic consonantal stems such as *fótr* 'feet', except for the *i*-umlaut. But whereas the *i*-umlaut alone differentiated the plural *fótr* from the singular *fótr*, this was not the case with *fjānd-r* : *fjānd-i*. The same principle holds true for the uncontracted forms *fīandr* : *fīandi*, and no doubt facilitated the borrowing of the vowel *a* in place of **e* in the plural. The rare plural form *friendr* (beside the usual form *frændr*) had no singular **friandi* to imitate; hence the vowel *e* was retained in spite of the combination *ia* in the verbal form *friða*.

3. **lR* > *ll*, BUT **lR* > *llr*. The assimilation of **lR* > *ll* regularly occurred only when the stem syllable (before the time of the syncope) was long or did not receive the chief stress; cf. **stōlaR* > **stōlR* > *stóll* 'chair', **hailaR* > **heilR* > *heill* 'whole'; **katilaR* > **ketilR* > *ketill* 'kettle'. On the other hand, **R* seems to have been exempt from this assimilation after *ll*, even though the double consonant rendered the stem syllable long; cf. **holpaR* > **holpR* > *holtr* 'gracious, favorable', **allaR* > **allR* > *allr* 'all'. Since the type *allr*, like the type *heill*, has a long stem syllable, it is plausible to assume that the **R* in the combination **lR* was likewise assimilated to **l*, yielding a phonetically correct form **all* (< **alll*) without the ending *-r*, just as in the type *heill*, and that therefore the *-r* in the type *all-r* is of secondary origin. This assumption is supported by the fact that in certain mss. the *r* (< **R*) is occasionally omitted; cf. gen. pl. *alla* (No. Hom.) for *allra*, gen. sg. fem. *allar* (St. Hom., *Plácítúsdrápa*, No. Hom.) for *allrar* (Noreen §277, 4a). Noreen is probably right in assuming that after secondary *ll* < **lp*, as in *holtr*, the *-r* is phonetically correct because **R* first fell together with original *r* after DENTAL CONSONANTS (§265), but his assumption that *allr* with original *ll* followed the pattern of *holtr* with original **lp* (§277, 4a) is in need of further clarification.

-a cannot, however, be correlated with the vowel *æ* in the *nd*-stem *frændr* pl. : *frændi* sg. 'kinsman', derived from the verb *friða* 'to love', since the origin of this *æ* is still problematical. According to the current view, the *æ* (through *i*-umlaut of **ā*) originally belonged only to the plural but was later transferred to the singular. But Sievers pointed out (*Beitr* 18.410) that the *æ* could originally have belonged also to the singular: 'PI **prē-ient* > PN **fræjand-* : **fræjind-*'.

⁴ Noreen §422 Anm. 1. Noreen assumes here that the vowel *a* in such Old Norwegian forms as *bóandr* 'peasants', *hafnandr* 'missbilliger' may be phonetically correct through lack of accentuation. But there is no reason to assume that in Old Icelandic *gefendr* < **gefand-iR* the accentuation was different from that in **fiendr* < **fiand-iR* (cf. *friendr*).

There is one factor which Noreen has overlooked in connection with this analogical *-r* in *all-r*, viz. that the ending *-r* represents a restoration of the phoneme *-r*, which was the sign of the masc. nom. sg. in the adjectival paradigm. It is therefore possible that the ending *-r* in the type *all-r* may be due to the pattern of the type *holl-r*, the *-r* acting as a case ending in conformity with the vast majority of adjectives in which the ending *-r* (< **-r*) was not lost through assimilation, such as *góðr* 'good', *ríkr* 'mighty', *skammr* 'short', *stórr* 'large'. Indeed, it is quite possible to account for this *-r* in certain words of the type *all-r* as a restoration of the *r*-ending without resorting to the analogy of the type *holl-r*. Take for example the verb *falla* (with original *ll*) 'to fall': **fall-iR* > **fellR* > **fell* > *fell-r* 'falls'; the ending *-r* can certainly represent simply a restoration of the usual verbal ending for the present 3rd pers. sg. indicative, as in the vast majority of verbs (cf. *ríð-r* 'rides', *byð-r* 'offers', *tek-r* 'takes', etc.). Occasionally, **lR* was assimilated to *ll* also in the verbal forms with SHORT stem syllables (because of the atonic position of the verb in the sentence); cf. **galir* > **gelR* > *gell* 'sings' (from *gala* 'to sing'). Alongside *gell* we find a later form *gel-r*, with *-r* restored. Since this *-r* represents the restoration of an ending which had been lost through assimilation, it follows that the *-r* in *fell-r* can similarly represent the restoration of an ending which had been lost through assimilation, especially since the *-r* in the present-tense form *fell-r* serves to differentiate this form from the 3rd sg. preterit *fell*. The *-r* in *fell-r* (< **fell*) may thus represent a restoration, not an analogical *-r* modeled on the phonetically correct *-r* in the combination *llr* (< *lþR*) in *holl-r*, though the result of the two processes is the same. In the adjectival type *all-r*, function alone does not seem sufficient to explain the restoration of the ending *-r*, since the *-r* was never restored in the adjectival type *heil-l*. It is significant, however, that where the initial *r*- in the adjectival endings *-ra* : *-rar* is omitted, viz. in the gen. pl. *all-a* (for *all-ra*) and the gen. sg. fem. *all-ar* (for *all-rar*), this *r* is not necessary as a sign of the case: the endings *-a* : *-ar* can serve this purpose, as in feminine substantives; cf. *all-a* : *axl-a*, *all-ar* : *axl-ar* from *qxl* 'shoulder'.

My conclusion then is that the restoration of the *r* after original *ll* was due to two analogical patterns: (1) the purely phonetic analogy of *-r* after secondary *ll* < **lþ* (*holl-r* : *all-r*), and (2) the analogical restoration of *-r* as a sign of the grammatical function (*gala*, *gell* > *gel-r* : *falla*, **fell* > *fell-r*). Since the two processes produced the same result, it is not possible to prove which was the point of departure; but from the foregoing analysis it is clear that both existed. The same hypothesis holds true for **nnR* > *nn* and **nþR* > *nnr*. Is the restoration of *-r* in the verbal form **brann-iR* > *brenn* > *brenn-r* due to the pattern of *gunn-r* 'war' (< **gunn-R* < **gunþ-R*) or to the need of an *-r* as a sign of the 3rd sg. present indicative of the verb *brenna* 'to burn'? Here Noreen again overlooks the grammatical function (§277, 4b).

4. THE RELATIVE-CONJUNCTIVE PARTICLE *es* (**ez* > **eR*) > *er*. Both forms existed side by side until about the year 1200, when the later *er* entirely superseded the original *es* (Noreen §473 Anm. 1). In spite of this, the *es* was at this time regularly retained (as *-s*) if it was attached as an enclitic particle, usually to

a pronoun or an adverb; cf. *sá er* : *sá-s* 'he who', *þar er* : *þar-s*, 'there where', *þegar er* : *þegar-s* 'as soon as'. (I have found only one exception to this rule, viz. *þá er* : *þá-r* : *þá-s* 'then when', where both forms survived.)

The question arises why the normalized secondary *er* should have been displaced by the original *es* as an enclitic particle. If *es* as an independent particle had already yielded *er*, we should expect the same to be true of the enclitic forms. The most plausible explanation is that in the enclitic forms the *-r* (< **-R*) could in certain examples either alter the preceding word through *R*-umlaut and thus disturb the correlation between the forms with and without enclitic, or through assimilation disappear as the sign of the relative-conjunctive function. If the *-r* was lost through phonetic change where the older phoneme *-s* was phonetically correct, there was no necessity for restoring the secondary *-r*. When the *-r* was not lost through phonetic change, it was displaced by the older *-s* because this *-s* was already attached to words in which the *-r* had been lost, i.e. the *s*-form had become extended and generalized. Examples illustrating the causes for the displacement of *-r* by *-s* are:

(1) *R*-umlaut: *sú-s* 'she who' : **sūr* > **sýr* (cf. *sýr* 'sow'); *nú-s* 'now when' : **nūr* > **nýr* (cf. *nýr* 'new'); *sá-s* 'he who' : *sār* > **sær*. The Runic form *sār* (Björketorp) existed before the time of the *R*-umlaut.

(2) The phoneme *-r* lost through assimilation but restored as *-s*: **stīðan-R* 'since' > *stīðan-n* (cf. **heiten-R* > *heitenn* 'called') : *stīðan-s*; **þann-R* 'him who' > *þann* (cf. **mann-R* > *mann* 'man') : *þann-s*; **hinn-R* 'he who' > *hinn* (cf. **kinn-R* > *kinn* 'chins') : *hinn-s*.

(3) The substitution of *-s* for phonetically correct *-r*: **þegar-R* 'as soon as' > **þegar-r* (cf. **hammar-R* > *hammarr* 'hammer') : *þegar-s*; **þar-R* 'there where' > **þar-r* (cf. **þur-R* > *þurr* 'dry') : *þar-s*.

5. *Halp* > *hjalp* 'HELPED'. The *j* in *hjalp* was obviously borrowed from the present *hjalpa*, where it was phonetically correct because of the *a*-breaking of the original stem vowel *e* (**helpa*). The point of departure for this analogy was the identity of the primary ablaut vowel *a* in *halp* with the secondary vowel *a* in the diphthong *ja* of *hjalpa*. In Old Icelandic, *hjalpa* was the only verb in the 3rd ablaut series with the breaking *ja* which shows this transference of *j* into the preterit singular. Later the process was extended to other verbs of this type (cf. Mod. Danish *gjælde*, *gjaldt* = ON *gjalda*, *galt* 'to pay'), and also to verbs in which a secondary *j* developed in the present tense (cf. Mod. Swedish *sjunga*, *sjöng*, but Danish *syng*, *sang* = OIcel. *syngva*, *song* 'to sing').

In the establishment and extension of this trend there may have been contributory factors not connected with the *j* of the present *hjalpa*. Such a factor could have been the strong verbs with a phonetically correct *j* before the stem vowel of the preterit singular but no diphthong *ja* due to *a*-breaking in the present—i.e. the reduplicating type *auka*, *jók* 'to increase', *hlaupa*, *hljóp* 'to run', etc. Here the *j* represents the survival of the vowel **e* (Goth. *at*) of the reduplicating prefix; it serves as a morpheme characteristic only of the preterit. The morpheme *j* resembled the secondary *j* of *hjalp*, and so could have contributed to its retention after it had been borrowed from the present, especially since this reduplicating class had a secondary preterit plural without *j* but with the same radical

vowel (u) as the plural *hulpum*; cf. *hljóp* : *hlupum* (for earlier *hljópum*), *hjalp* : *hulpum*. So far as the *j* is concerned, the secondary form *hjalp* had converged with the type *hljóp*, i.e. it had entered into a new pattern, which contributed to its retention and later extension.

6. *Var* > *vár* 'WAS'. Noreen (§498 Anm. 7) rightly explains the lengthening of the vowel by the influence of the pattern *át* : *ótum* 'ate', in which the singular *át* was phonetically correct, parallel to the late anomalous forms *báð* (< *bað*) 'asked', *sát* (< *sat*) 'sat', etc.⁵ He evidently overlooked, however, the fact that in §126.1 he had already explained the lengthening of the vowel in *vár* as a phonetic process, due to the influence of the following *r* < **R* in such forms as *úr ór* 'out of' (< **ūr *ōR* < **uz* = Goth. *us*) and *ér* 'is' (< **ēR* < **eR* < **iz*).⁶ This latter explanation cannot be correct, since the *-r* in *va-r* was borrowed from the plural paradigm (*va-r* : *vó-r-um*) in place of Proto-Germanic *-s* (*vas* : *vórum* = Goth. *was* : *wēsūm*). At the time of the borrowing, the PN **R* of the plural paradigm had already fallen together with original *r*; but when this was transferred to the singular, it could not affect the length of the preceding vowel. On the other hand, the lengthening of the vowel in the verbal form *ér*, which Noreen assumes to be parallel to the lengthening of the vowel in *vár* (fn. 6), is phonetically correct. Proto-Germanic **ist* (= Goth. and OHG *ist*) appears in the Runic inscriptions (Amle, Möjebro) as *iR*, from which we may postulate a PN form **iz*; hence **iz* > *iR* > **eR* > **ēR* > *ér* > *er* in atonic position.

7. *Sekkja*, *sekþa* : *sekta*. Since the stem syllable (**sak-*) of the *jan*-verb *sekk-ja* (< **sak-jan*) 'to declare guilty; inflict punishment' is short, we should expect the preterit forms to be without *i*-umlaut, viz. **sak-þa* : **sak-ta* on a level with *vekkja*, *vak-þa* : *vak-ta* 'to awaken'. *Sekkja* is a denominative verb derived from the adjective *sek(k)r* 'guilty'. Axel Kock⁷ explains this irregular *i*-umlaut in *sekþa* : *sekta* on the ground that these forms do not represent the preterit of *sekkja* but of a denominative verb **sekta* derived from the substantive *sekt* 'guilt, punishment'. He bases this assumption on the East Norse forms, *sækta* (inf.) : *sækta* (pret.).

This assumption, however, is fundamentally weak: it gives no reason why the preterit forms *sekþa* : *sekta* should have been borrowed from a verb **sekta*. The long-stem preterit *sekta* (inf. **sekta*) could just as well have been re-formed to **sakta* after the pattern of the phonetically correct short-stem form **sakta* (inf. *sekkja*). All that the East Norse preterit *sækta* shows is that it is the phonetic equivalent of West Norse *sekta* and therefore furnishes no evidence that the two forms go back to the same infinitive form, especially since a West Norse infinitive form **sekta* is not recorded.

It is evident that the forms *sekþa* : *sekta* owe their *i*-umlaut to the influence of the *jan*-verbs with long stem syllables; but in order to show this influence we must discover why it was operative. To this end we postulate a point of de-

⁵ Cf. §498 Anm. 7: '*báð ... sát ... vár ... zu bǫþom usw. neugebildet (nach át : ótum ...)*'.

⁶ Cf. §126.1 under *Vokaldehnung*: 'Vor tautosyllabischem, aus urn. *R* (urgerm. *z*) entstandenem *r* ... präpos. *ór*, *úr* ... präp. *vár* (No. Hom. 8 mal ... gew. *var* weil unbetont) war ... und *ér* (gew. *er*) verb'.

⁷ *Beiträge* 18.436.

parture in which the two types of *jan*-verbs, with long and with short syllables, have peculiar phonetic characteristics in common, not one in which they are different, as Kock assumes (cf. **sakta* : *sekta*, the result of the analogy). Such a point of departure is furnished by the infinitive of both types with *i*-umlaut of the radical vowel **a* > *e*, in which the stem syllable ended in a geminate *kk*. In the short-stem type the geminate is secondary, the result of the following *j*-suffix, as in **sak-jan* > *sek-ja* > *sekk-ja*. In the long-stem type the geminate is primary, resulting from the assimilation of **nk*, as in **bank-jan* > *bekk-ja*. There are two verbs *bekkja*, a primary one meaning 'to be aware, observe' (pret. *þátta* = Goth. *þáhta*), and a secondary one, derived from the adjective *pekk* 'pleasant', meaning 'to recognize, think well of'. The preterits *pekþa* : *pekta* of secondary *bekkja* follow the regular pattern of the *jan*-verbs. It is therefore reasonable to assume that the verb *sekk-ja*, with an original short stem syllable, followed the pattern of the type *bekk-ja* with a long stem syllable, discarding the phonetically correct preterits **sakþa* : **sakta* in favor of *sekþa* : *sekta* after *pekþa* : *pekta*. That this leveling did not occur with other verbs of the type *sekk-ja*, such as *vekk-ja*, may be due to the fact that both *sekk-ja* and *bekk-ja* are derived from adjectives which contain the *j*-umlaut of the radical vowel.

COGNITION IN ETHNOLINGUISTICS

ERIC H. LENNEBERG

Massachusetts Institute of Technology

The republication of Benjamin L. Whorf's articles on what Trager calls metalinguistics has aroused a new interest in this country in the problem of the relationship that a particular language may have to its speakers' cognitive processes. Does the structure of a given language affect the thoughts (or thought potential), the memory, the perception, the learning ability of those who speak that language? These questions have often been asked and many attempts have been made to answer them.¹ The present paper is an attempt to lay bare the logical structure of this type of investigation.

CRITICAL RETROSPECT

A basic assumption. Underlying all of Whorf's theoretical work is the fundamental assumption that the individual's conception of the world (including perception, abstraction, rationalization, categorization) is intimately related to the nature of his native language.² Throughout his work Whorf illustrates this idea with examples from American Indian languages, showing how they differ from English. However, a demonstration that certain languages differ from each other suggests but does not prove that the speakers of these languages differ from each other as a group in their psychological potentialities. To prove this, it would be necessary to show first that certain aspects of language have a direct influence on or connection with a given psychological mechanism, or at least that speakers of different languages differ along certain psychological parameters. In addition to comparative data Whorf adduces occasionally a different type of evidence. An example is his analysis of many hundreds of reports of circumstances sur-

¹ Bibliographies of the voluminous literature may be found in the following works: Kurt Goldstein, *Language and language disturbances* (New York, 1948); Friedrich Kainz, *Psychologie der Sprache* (Stuttgart, 1941/43); George A. Miller, *Language and communication* (New York, 1952); Charles Morris, *Signs, language and behavior* (New York, 1946); David L. Olmsted, *Ethnolinguistics so far* (*SIL*, Occasional papers, No. 2; 1950); N. H. Pronko, Language and psycholinguistics: A review, *Psych. bull.* 43.189-239 (1946).

This paper was stimulated by research carried on under the auspices of the Values Study in the Laboratory of Social Relation, Harvard University, and the Communications Project at the Center for International Studies, Massachusetts Institute of Technology. I wish to express my thanks to both institutions. I am also greatly indebted to Harry Hoijer for inviting me to participate in the Conference on Ethnolinguistics, held in Chicago during March 1953, where the discussion of some of the problems raised in this paper helped to clarify my thoughts. Finally I gratefully acknowledge the many helpful suggestions made to me by Noam Chomsky, who read two earlier versions of this article.

² Whorf is not alone in making this assumption. Cf. Dorothy D. Lee, Linguistic reflection of Wintu thought, *IJAL* 10.181-7 (1944); Lucien Lévy-Bruhl, *Les fonctions mentales dans les sociétés inférieures*, Ch. 4 (Paris, 1910); Leo Weisgerber, Adjektivistische und verbale Auffassung der Gesichtsempfindungen, *Wörter und Sachen* 12.197-226 (1929). The last of these is a representative of what H. Basilius has called Neo-Humboldtian ethnolinguistics, *Word* 8.95-105 (1952); the entire movement is based on the assumption discussed here.

rounding the start of fires, for instance the empty-gas-drum case.³ An explosion had been caused by an individual who had carelessly flung a burning cigarette stub into a gas drum which this person in his insurance report called *empty*. Whorf argues that the individual's carelessness was caused by the fact that the word *empty* has two different meanings in English: (1) null and void, negative, inert, and (2) a space which may contain nothing but a vapor, liquid vestiges, or stray rubbish. The English language forced the individual to call the gas drum *empty*, and think of it in terms of that word. Since this word could mean null and void, Whorf argues that the presence of explosive vapors and inflammable liquid vestiges could be disregarded by the speaker, who then behaved towards the drum as if it were absolutely empty. I cannot accept this as evidence for the assumption that behavior is influenced by language. Clearly, English is capable of distinguishing between a drum filled with an explosive vapor, one that contains only air, and one which is void of any matter. This very sentence is my evidence. The person who caused the fire could have replaced the word *empty* by *filled with explosive vapor*. His failing to do so (as well as his careless behavior) points to a lack of experience with explosive vapors, perhaps complete ignorance of their existence. The linguistic—or rather stylistic—fact of the occurrence of the word *empty* in the individual's insurance report would indeed be interesting if Whorf could have shown at the same time that this man had had plenty of contact with and knowledge of the explosive vapors which form in emptied gas drums. This Whorf did not try to do. In short, the basic assumption that language affects non-linguistic behavior derives from an inspection of linguistic facts. Therefore nothing is added to such an hypothesis by referring back to the same or similar linguistic facts.

Translation. (a) Translation, while useful for the formulation of working hypotheses of the most exploratory nature, is in itself an inadequate way towards the finding of objective facts. Obvious as this may seem, it is necessary to spell out in detail the shortcomings of the translation method in ethnolinguistics.

(b) I illustrate my point with another example taken from Whorf.⁴ After posing the question: 'What do different languages do ... with the flowing face of nature ...?', Whorf answers: 'Here we find differences in segmentation and selection of basic terms. We might isolate something in nature by saying, "It is a dripping spring." Apache erects the statement on a verb *ga*: "be white (including clear, uncolored, and so on)." With the prefix *no*-, the meaning of downward motion enters: "whiteness moves downward." Then *to*, meaning both "water" and "spring" is prefixed. The result corresponds to our "dripping spring," but synthetically it is: "as water, or springs, whiteness moves downward." How utterly unlike our way of thinking! [NB!] The same verb, *ga*, with a prefix that means "a place manifests the condition" becomes *gohlga*: "the place is white, clear; a clearing, a plain." These examples show that some languages have means of expression ... in which the separate terms are not as separate as in English but flow together into plastic synthetic creations.' Whorf analyzes the Apachean

³ Whorf, The relation of habitual thought and behavior to language, *Language, culture, and personality* 75-93 (Menasha, Wis., 1941).

⁴ Whorf, Languages and logic, *The technology review*, Vol. 43 (1941).

statement by giving the English equivalent for the general meaning of each Apachean element, and then compares the resulting sequence of meanings to the phrase, 'it is a dripping spring'. The sequence of meanings (i.e. the glosses) and the English phrase are not, however, quite comparable. Whorf does not give the general meaning of the English morphemes. If he had, something like this would have resulted:

it, any object or organism which is not an adult human being; *is*, particle which denotes that what follows is a predicate of what precedes; *a*, particle which denotes that what follows is to be understood generically, not specifically; *drip(p)*, process in which any liquid falls in small natural segments; *-ing*, particle which denotes that the preceding process has not come to an end; *spring*, something that is not static (eruption of water, device to make mattresses elastic, and so on).

To abstract a general meaning of a morpheme or lexeme may occasionally be of some methodological use; but we must not confuse such an abstraction with an isolable segment of an utterance. General meanings lack reality, so to speak. It makes no sense to equate the global meaning of an utterance with the sequence of abstracted, general meanings of the morphemes that occur in that utterance. To translate the Apachean statement *it is a dripping spring* appears no less reasonable than to translate it *as water or springs, whiteness moves downward at a place* (or, *the place is white, clear; a clearing; a plain*—which, I gather from Whorf, is the synthesis of the elements); for what we translate are equivalent verbal responses to particular stimulus situations, and the Apachean response to the natural phenomenon in question corresponds to our response *it is a dripping spring*. This type of linguistic evidence, therefore, stands or falls with our philosophy of translation. It might be objected here that Whorf's evidence is not the translation itself but the fact that the Apachean's verbal response to this natural phenomenon is the same as his verbal response to a different phenomenon, namely one to which we respond *the place is white, clear, etc.* and that the Apachean therefore makes a single response to stimuli to which we make distinct responses. This objection, while touching upon an important problem, does not justify the translation method. For what we really want to know is how the Apachean structure of syntactic categories differs from the English one. Translation cannot answer this problem. Through it—and that is its value—we merely know that the problem is not a spurious one.

(c) A further objection to translation as a sufficient method in this type of research is that it actually vitiates the attempt to demonstrate cognitive difference as evidenced in two or more languages. For, if a language were actually an aspect of a particular psychological make-up or state of mind (or more precisely, an aspect of a cognitive process, which is not to be confused with the thought content), then, in the process of translation, we would be substituting the psychological elements characteristic of one make-up for those of another, so that we would finally compare two sets of elements of one and the same psychological structure.

(d) There is a metaphorical element in language per se. The literal meaning of many metaphors, especially the most frequent ones, never penetrates conscious-

ness, e.g. *everybody, in the face of, beforehand, breakfast, inside, already*. The translation method, however, distorts the significance of such forms of speech and often induces investigators to draw rather ludicrous conclusions. To illustrate the mentality of certain African tribes, Cassirer writes:⁵ 'The languages of the Sudan usually express the circumstance that a subject is in process of action by means of a locution which really means [NB!] that the subject is *inside* that action. But since, moreover, this *inside* is usually expressed very concretely, phrases result such as *I am on the inside of walking, I am the belly of walking*, for "I am in the process of walking".'

(e) The process of taking stock of general meanings, which underlies translations, engenders the belief that languages can convey no more and no less than the general meanings of morphemes. It seems more fruitful to assume⁶ that much more is cognized than is expressed by individual morphemes. Morphemes and their meaning are regarded more appropriately as mnemotechnical pegs of a whole situation which is brought into consciousness by the statement as a whole. The general meaning of morphemes is probably of lesser importance in cognition than the SUM OF ASSOCIATIONS bound up with the complete utterance, or even with individual morphemes or groups of morphemes.

(f) When the translation involves a juxtaposition of totally different cultures (say Chukchee and English) we are not only faced with a semantic problem. No matter what precautions we take in glossing a word, almost no correspondences can be established between many denotata. For instance, the cultural and physical contexts of Chukchee utterances are, with a few exceptions, incomparable with the contexts within which English is spoken. Chukchee weapons, food, manners, standards of any sort, landscape, fauna, and flora are mostly unfamiliar to English-speaking cultures. Thus, practically no common frame of reference, no basis for a segmental, one-by-one comparison exists between these two languages. Translation here can be only a very rough approximation of what has been said and intended originally.

Ad-hoc theories. It is a commonplace in scientific methodology to avoid etiological theories which are incapable of satisfying more than one single and specific occurrence of events; yet by necessity working hypotheses often have to be of this nature. We see a picture fall off a wall directly after hearing a dog bark in the neighborhood. As a working hypothesis the two events might be causally related. Upon verification of the hypothesis we note, however, that in general barking is not followed by things dropping to the ground, nor is the falling of pictures from the wall usually preceded by barking or similar noises. We are unable in this instance to formulate a theory because the working hypothesis cannot be generalized. Turning to ethnolinguistic literature we find an abundance of working hypotheses where it is difficult to see how they might contribute to a universally valid and useful theory of language (such that language is related

⁵ Ernst Cassirer, *Philosophie der symbolischen Formen: Die Sprache* 1.168-9 (Berlin, 1923). Cassirer's source is D. Westermann, *Sudansprachen*.

⁶ This assertion and the following are based on evidence from experimental psychology. See George Humphrey, *Thinking: An introduction to experimental psychology*, Chs. 4 and 8 (London, 1951); Miller, *Language and communication* passim.

to non-linguistic behavior), because the facts underlying such working hypotheses cannot be generalized so as to fit more than a single language.⁷ I am not saying that such hypotheses are right or wrong; many have been proposed by experts on specific cultures, by scientists of undisputed merit. I am merely pointing to the difficulty, if not impossibility, of deducing from these hypotheses, if they are sound, general and verifiable laws. A common means of validating hypotheses has been barred from the beginning in these cases, namely cross-cultural verification. This, however, does not exclude the possibility that the investigators may have intra-cultural evidence for each individual hypothesis proposed.

TOWARDS A METHODOLOGY

Codification and cognition. (a) A basic maxim in linguistics is that anything can be expressed in any language.⁸ There may be differences in the ease and facility for the expression of certain things among various languages but at present we do not know whether this difference in ease is attributable to the properties of a given language qua vehicle of communication⁹ or to the cultural development of the speakers. In fact, this is one of the problems to be solved in ethnolinguistics. Now, if we believe, as we do, that we CAN say anything we wish in any language, then it would seem as if the content or subject matter of utterances does not characterize or, indeed, give us any clear information on the communicative properties of a language. Thus we are led to the somewhat banal conclusion that the only pertinent linguistic data in this type of research is the HOW of communication and not the WHAT. This HOW I call the codification; the WHAT I call the messages. Codification can be studied in three phases: (1) the process of encoding; (2) the code; (3) the process of decoding.¹⁰ The study of the code results, for instance, in statements about the structure of phonemes, morphemes, and syntactic categories; about acoustic characteristics of speech sounds; about the frequency distributions and the transitional probabilities of given segments; about the efficiency of the code within stated contexts. In these instances meaning can be excluded entirely from our research, at least theoretically, and we have therefore an assurance that we are actually studying aspects of codification. Unfortunately, however, it is not always equally easy to decide whether a phenomenon is pertinent to codification or not. Many assertions about language which derive from semantic observations or, at any rate, which include elements of meaning, nevertheless seem to be relevant to codification. Most

⁷ Most of Whorf's and Dorthy Lee's working hypotheses are of this nature. Harry Hoijer's Cultural implications of some Navaho linguistic categories, *Lg.* 27.111-20 (1951), and the tentative connections between various linguistic features and nonlinguistic behavior mentioned by Claude Lévy-Strauss, *Language and the analysis of social laws*, *Amer. anthr.* 53.155-63 (1951), also fall into this category of working hypotheses.

⁸ Cf. Sapir, The grammarian and his language, *Selected writings of Edward Sapir* 153-4 (Berkeley, 1949). It is assumed here that any vocabulary can be expanded.

⁹ The use of the term *vehicle of communication* does not mean that I deny (or even take a position toward) the epistemological contention that language and knowledge are indistinguishable. I am merely referring to the communicative capacities of language.

¹⁰ John B. Carroll, *Report and recommendations of the Inter-disciplinary Summer Seminar in Psychology and Linguistics* 8 (Ithaca, N. Y., 1951).

obvious in this connection is the fact that a language always selects for codification highly specific aspects from the physical and social environment. This raises two questions: How can we describe objectively the aspects that are being selected out of a great number of other possible aspects? Why are these aspects selected and not others? There can be little doubt that these considerations, though clearly of a semantic character, have a bearing on the problem of codification. Hence, the distinction between codification and messages is not the same as between syntactics and semantics or between form and meaning. All those observations about meaning are relevant to codification which refer to an aspect of speech behavior which is forced upon the individual speaker by the rules of his language and where infringement of the rules would result in defective communication. For instance, an individual reporting about a given event is forced to stipulate very definite conditions, aspects, and relationships if he wants to be understood.¹¹ However, he is free to report on the event in the first place, and also to elaborate on circumstances of the event which are not included in the compulsory stipulations. Whatever information is optional in his communication is message.

(b) Once we have clearly isolated data on codification, such data may be related hypothetically to nonlinguistic behavior. If the researcher is interested in cognition,¹² as I am, he will investigate relations that obtain between codification and such behavior as is indicative of memory, recognition, learning, problem solving, concept formation, and perception, hoping to show that certain peculiarities in these processes can be explained by—and only by—knowledge of the speakers' peculiarities of codification.¹³

The intra-cultural approach. (a) Ethnolinguistic research based on cross-cultural comparison must endeavor to isolate data, both on codification and on cognition, that are general enough to have comparable equivalents in at least two different languages and cultures; otherwise comparison would be meaningless. It is not infrequent, however, that a working hypothesis relates a certain cognitive datum to some phenomenon pertinent to codification which appears to be unique, lacking entirely a parallel in any other language. There is a simple way of studying this situation; I call it the intra-cultural approach, because it reduces cross-cultural comparison to a desirable but not indispensable expansion of investigations. This method is so easy to manipulate that many investigators may perhaps come to use it even where the cross-cultural approach is applicable directly.

(b) I begin with a practical demonstration of the method. Problem: Languages differ in their systems of classifying the ten million odd colors which every normal individual can discriminate.¹⁴ Under laboratory conditions the power of color

¹¹ These conditions, aspects, and relationships are primarily but not exclusively expressed by grammatical categories.

¹² For a modern definition of this term see Robert Leeper, *Cognitive processes, Handbook of experimental psychology* 730-57 (ed. S. S. Stevens; New York, 1951).

¹³ What I am proposing to do here is not in principle different from what Whorf (for instance) occasionally suggested. The difference between Whorf and me is rather in our respective attempts to substantiate our hypotheses.

¹⁴ Cf. Ralph M. Evans, *An introduction to color* 230 (New York, 1948).

discrimination is probably the same for all human beings, irrespective of their language background. But we do not know whether the habitual grouping of colors, according to certain labels provided by every language, might not affect some other cognitive processes involving color stimuli. To be more specific, in English obviously not all colors are named with equal ease and unambiguity. Do English-speaking people therefore recognize easily-named (i.e. highly codeable) colors with greater facility than colors not so easily named?¹⁵

The first step toward solution of this problem is to ascertain the linguistic facts.¹⁶ A representative sample of English speakers is drawn and a number of colors are prepared that have comparable perceptual properties. Then the notion 'codeable' is investigated and defined operationally, so that we can divide the physical color stimuli by means of one or a combination of a few simple criteria, into two groups: one consisting of 'highly codeable' and one of 'less codeable' colors. I must omit here the details of this procedure and also the reasoning that underlies the individual steps leading to the development of such a criterion. Let me simply state that UNANIMITY IN RESPONSE proves to be a useful criterion (among others) in this connection. Some colors are consistently given the same name by every speaker; others are given a variety of names, sometimes as many names as there are subjects. Regarding the speakers now as a group giving a linguistic response to each color, we may say that some colors have the property of eliciting a homogeneous response from English-speakers, whereas other colors elicit a heterogeneous response. This is to say that linguistic communication in English is more efficient when some colors are referred to than when others are.¹⁷ There are cogent reasons to assume that the distinction made here between the colors is a purely linguistic one, and that there are no physical properties in the colors or physiological ones in the eye which would elucidate the difference in response made by English-speakers to these colors.¹⁸

¹⁵ This is a specific question within a problem that has been posed by many other investigators. Sapir said: 'Language is a ... self-contained, creative symbolic organization, which not only refers to experience largely acquired without its help, but actually defines experience for us by reason of its formal completeness and because of our unconscious projection of its implicit expectations into the field of experience.' (Conceptual categories in primitive languages, quoted by I. J. Lee, *The language of wisdom and folly* 265 [New York, 1949].) Sapir makes the same point in *The status of linguistics as a science, Selected writings* 162.

¹⁶ The following is an outline of research in progress carried on by Roger Brown of Harvard University and myself. The details of the project will be published as soon as the data are fully assembled.

¹⁷ If there is no well defined name for a color, it is reasonable to assume that linguistic communication about it is poor.

¹⁸ Again space does not permit me to cite all the evidence in support of this assertion. The interested reader may inspect the colors used; they are produced by the Munsell Color Co., a scientific research organization. Most of them are published in the two volumes of the *Munsell book of colors* (Baltimore, 1921 and 1942). Codeable colors have the notation 2.5 PB/7/6, 5 PB/4/10, 10 P/3/10, 5 RP/6/10, 5 YR/3/4, 3 GY/7.5/11.2, 7.5 GY/3/4, 2.5 G/5/8, 5 Y/8/12, 7.5 G/8/4, 2.5 R/7/8. Non-codeable colors have the notation 10 BG/6/6, 8.5 B/3/6.8, 10 PB/5/10, 2.5 R/5/10, 8 RP/3.4/12.1, 7.5 R/8/4, 2.5 Y/7/10, 7.5 Y/6/8, 7.5 YR/5/8, 5 P/8/4, 5 BG/3/6. Colorimetric and psycho-physical data on these colors are published in *Journal of the Optical Society of America* 30.573-645.

The next step is to determine whether there is a difference in ease of recognition by English-speakers between the colors constituting one group and those constituting the other. The two groups, I repeat, are perfectly balanced in physical and perceptual properties; the only difference is that the colors in one group have well defined names in English, whereas the colors in the other do not.¹⁹ If we now use, in random order, colors belonging to either group, say ten from each, in a standard recognition test,²⁰ we can easily discover whether English speakers do better when they have to recognize colors which are highly codeable in their language than when they recognize less codeable colors. In the actual performance of the experiment²¹ this appears to be the case. Statistically, codeable colors are recognized significantly more often than less codeable ones, and thus there is good evidence that the particular linguistic fact, codeability, affects the cognitive process, recognition.

(c) Suppose now that this entire color research were repeated in a different culture where a different language is spoken. If our predictions about recognition, based on previously determined facts of codification (which vary of course from language to language), should not be borne out in this other language, the argument advanced in the first experiment would be seriously weakened. Conversely, if the results should be confirmed, this would fortify the argument. In either case, however, VALIDATION of the basic hypothesis is independent of cross-cultural comparison. The cross-cultural comparison merely adds or subtracts weight. It is very important to realize that the validation itself is the result of intra-cultural correlation of two sets of recognition behavior on the one hand (in the described context we may say 'good' and 'bad' behavior) with two sets of English speech behavior on the other hand (efficient and not so efficient linguistic communication). It appears that recognition behavior is inefficient where speech behavior is inefficient.

(d) Not only is the validity of this experiment independent of cross-cultural comparison; but if cross-cultural comparison is desired, the method dispenses with the necessity for translation, or the exact equation of linguistic data between one language and another. For what will be compared are CORRELATIONS OF SPEECH BEHAVIOR WITH RECOGNITION BEHAVIOR, not linguistic forms. Superficially it may look as if the translation method were implicitly the same as the intra-cultural method, for both methods seem to be concerned with the meaning of certain linguistic forms which are being compared. However, the intra-cultural method resembles the translation method only in its very elementary and primary step: both methods recognize the existence of a problem on the grounds of intui-

¹⁹ Codeability of colors does not seem to be linked to cultural importance or preference for these colors. The reader may convince himself of this by trying to name all the colors in his environment. He will notice that colors for which he has a 'good' name occur much less frequently than colors which are difficult to label unambiguously.

²⁰ Such tests are described in K. Koffka, *Principles of Gestalt psychology*, Chs. 11-3 (New York, 1935).

²¹ The test colors were exposed four at a time, for two seconds. After a waiting period of thirty seconds, subjects had to find the test colors on a color chart of 120 colors. All colors were identified by numbers. The subjects used in this experiment were not required to use any color name whatever.

tive knowledge of the meaning of forms. The translation method defines meanings by trying to equate forms of a language foreign to the investigator to forms of his native language (where meanings are said to be known). The intra-cultural method need not rely on this haphazard procedure; instead, it objectifies the intuited meanings of forms by carefully relating them to stimuli of the environment. Thus it is possible (at least in some instances) to specify meaning by referral to the physical properties of those stimuli.

(e) Stated in general terms, the intra-cultural approach consists of the following. Some aspect of codification is described in order to correlate it with non-linguistic behavior. A frame of reference is established in terms of which both the speech behavior and the non-linguistic behavior can be described or specified; a particularly convenient frame of reference is the physical environment within which both types of behavior take place. In the experiment described, the frame of reference was provided by the stimuli sensed as colors. The speech events (color terms) and the behavioral events (recognition) were related to these stimuli. The specifications of the physical properties of the stimuli served as a metalanguage, so to speak, for the description of both types of events.

The fundamental principle of the intra-cultural method is that the physical stimuli, whatever they may be, can be classified on the grounds of linguistic criteria so that the constituents of each class are all characterized by the particular way in which they are codified. It is necessary that the codification criterion should be the *ONLY* criterion by which the stimuli can be grouped in this way. If now the non-linguistic behavior in response to the stimuli thus classified varies systematically in accordance with the class to which the individual stimulus has been assigned, we may attribute such regular variation in non-linguistic behavior to the regular variation in the speech correlates.

REVIEWS

Preliminaries to speech analysis: The distinctive features and their correlates. By ROMAN JAKOBSON, C. GUNNAR M. FANT, and MORRIS HALLE. (Technical report No. 13, 2d printing with additions and corrections.) Pp. viii, 58. [Cambridge, Mass.]: Acoustics Laboratory, Massachusetts Institute of Technology, 1952.

Reviewed by PAUL L. GARVIN, *Institute of Languages and Linguistics*

If, to use Householder's ingenious labeling,¹ there have been quite a few important theoretical statements of 'hocuspocus' linguistics lately, Jakobson and his collaborators have definitely contributed to the 'God's truth' side of the ledger. This is, however, not the only linguistic issue of the day on which the *Preliminaries* represent a definite and controversial viewpoint. There are, in my opinion, four major areas of disagreement among linguists where the contribution of Jakobson, Fant, and Halle takes sides, either implicitly or explicitly.

(1) The status of linguistics as a discipline: is it primarily a set of operations serving to organize into a structure a set of otherwise disparate data ('hocuspocus' linguistics), or is it a cognitive science, whose aim it is to gain knowledge of an existing structure which manifests itself in the data ('God's truth' linguistics)?

(2) The status of meaning in linguistics: is meaning part of the linguistic data, or does it belong in another set of phenomena—metalinguistics, if you like,² to be used in linguistic operations only as a 'shortcut' in lieu of more valid procedures independent of meaning?³

(3) The status of phonetic data in linguistics: are they linguistic in nature, to be used in the statement (or discovery) of phonemic patterns, or should phonemic analysis be based on distribution alone, with phonetic data considered mere irrelevant substance?⁴

(4) The status of linguistic contrasts: are linguistic patterns based on (or statable in terms of) binary contrasts, or is it a tour de force to reduce them to such symmetrical terms?⁵

The *Preliminaries* take the well-known Jakobsonian position on all four of these questions: (1) linguistics is a cognitive science; (2) meaning is an integral part of language and hence cannot be eliminated from linguistic statements; (3) phonetic facts have important bearing on phonemic structures; (4) phonemic

¹ Fred W. Householder Jr., review of Harris' *Methods in structural linguistics*, *IJAL* 18.260 (1952).

² George L. Trager, *The field of linguistics* (*SIL*, Occasional papers, No. 1; Norman, Okla., 1949).

³ Bernard Bloch, A set of postulates for phonemic analysis, *Lg.* 24.5 (1948).

⁴ Louis Hjelmslev (transl. by Francis J. Whitfield), *Prolegomena to a theory of language* 48 ff. (Indiana University Publications in Anthropology and Linguistics, No. 7; 1953).

⁵ Cf. Zellig S. Harris' criticism of neutralization, *Simultaneous components in phonology*, *Lg.* 20.183 (1944).

structures (the *Preliminaries* are not concerned with morphemes) pattern in terms of binary contrasts.

It is worth underlining here that the Jakobsonian position on these issues implies no neglect either of distributional facts or of operational procedures. On the contrary: on the one hand distribution accounts for some of the crucial points in Jakobson-Fant-Halle's argument, on the other hand distributional statements are extended beyond the type of data usually covered by them. Operational procedures are, in Jakobsonian terms, not so much classifying devices as means of discovering and describing the reality of structure.

1. THE STATUS OF LINGUISTICS AS A DISCIPLINE. The Jakobsonian metaphysics appears quite clearly from one of the earliest statements (1) in the *Preliminaries*. Given a certain sample of speech (a corpus), 'The question arises: How many significant units ... do the sound-shapes of the sample contain?' These are elements of the existing structure for the linguist to FIND, not elements of the description to be introduced by the linguist for the convenience of his analysis.⁶ This assumption of the reality of language as a system underlying the observable speech behavior recorded in the corpus is clearly based on the Saussurian distinction between *langue* and *parole*;⁷ it is opposed to the assumption that the corpus with its recurrent partials, being an observable reality, is also the only reality. The authors' choice of the term *sample* for their corpus is not, I believe, an accidental one: they do not describe their corpus as the ultimate object of their study,⁸ but merely as a sample of the underlying system which is the real object of linguistic cognition.

A number of logical consequences (call them corollaries) follow from this assumption. If there is such an object as *la langue*, then a description of this object must be a statement about the properties setting it off from other, comparable objects, and at the same time allowing it to be included with these same objects in a larger class. The latter was accomplished by de Saussure in defining language as one of many existing systems of signs, and including linguistics in a larger science of semiology.⁹ It means that language shares some fundamental characteristics with other sign systems; this allows Jakobson-Fant-Halle to draw upon some of the basic insights of communication theory to throw light upon sign systems in general (17 and *passim*). It also means that some basic assumptions have to be made concerning the nature of language as a system which govern operational procedures and have to be adhered to throughout the analysis. In a sense, therefore, the Jakobsonian answer to the other three questions follows logically from the answer to the first.

Two objections are most often made to the assumption that a linguistic system

⁶ Thus Charles F. Hockett, Componential analysis of Sierra Popoluca, *IJAL* 13.259 (1947): 'shall we operate [NB] in terms of recurrent motions and positions of the speech organs, or in terms of sound-features in the acoustic sense?'

⁷ See Rulon S. Wells, De Saussure's system of linguistics, *Word* 3.1-31 (1947), esp. 15-8.

⁸ Cf. Hockett, A formal statement of morpheme analysis, *SIL* 10.29 (1952): 'we mean that we seek the easiest way to describe all the utterances of our corpus, without implying anything about possible utterances in the language which do not occur in the corpus.'

⁹ Ferdinand de Saussure, *Cours de linguistique générale* 33-4 (Paris, 1949).

exists which is not merely a product of the linguist's analytical procedures. One is epistemological: how do we know that there is an underlying pattern, since the corpus is the only observable reality, and any statements beyond this observable reality can be only inferential? The second I might call operational: do we not impose undue limitations on our procedures by making a-priori assumptions? Why not just limit ourselves to the observable facts of the corpus and procede from there? The best answer to both objections is in my opinion given by Hjelmslev, who has elaborated the Saussurian concept of *langue* and *parole* into a complex logical system. It is an operational answer to both the epistemological and the operational objection against the distinction *langue-parole*: we MUST assume the existence of a system underlying the corpus (as Hjelmslev puts it, *op.cit.* 24, the process [corpus] determines [presupposes] the system) because there is no other way of reducing the potentially infinite corpus to a finite number of elements composing a system: 'if no restricted inventory appeared however long the analysis were continued, an exhaustive description would be impossible' (26; cf. 9-10).

The problem of the epistemological status of the underlying structure is of course not limited to linguistics. It exists both in social anthropology and in social psychology.¹⁰ For the latter field, T. M. Newcomb formulates the relationship of system to observable behavior in regard to the crucial problem of social position and role as follows:¹¹ 'A role, as I think the term should be used by social psychologists, thus consists of a limited set of behaviors "tied together" by a common understanding of the functions of a position. Not all of these behaviors will be performed by every occupant of the position, nor will they be performed in the same way by any two persons as they take the same role. The actual behavior of any individual in taking a role must be sharply distinguished [NB] from the role itself. Every role is necessarily defined in relation to one or more other roles ... The meaning of any role is thus to be seen in the entire network of roles, or role system, of which it is part.' The role system of which Newcomb speaks, and the behavior in taking a role, clearly parallel the *langue-parole* distinction; Newcomb speaks of role behavior as being PERFORMED, just as Jakobson-Fant-Halle speak (5) of phonemic oppositions being IMPLEMENTED phonetically.

Newcomb's conceptual framework is useful in understanding still another aspect of Jakobsonian theory. Role systems are defined in terms of social norms, a conception which is gaining increasing acceptance among social psychologists. Social norms, in turn, have to be formulated (*op.cit.* 35) 'in terms of perceptual processes', in terms of 'shared ways of perceiving things'; in social psychological terms it is quite legitimate to refer to language as a system of social norms, and hence also to interpret it in terms of shared frameworks of perception. This is precisely what Jakobson-Fant-Halle do (10): 'it is ... the whole patterning of

¹⁰ Note incidentally the Saussurian position of placing linguistics in 'semiology', which in turn is a part of 'sociology'. (The latter, to the French, includes both social anthropology and social psychology.)

¹¹ Theodore M. Newcomb, *Social psychological theory, Social psychology at the crossroads* 36 (ed. John M. Rohrer and Muzafer Sherif; New York, 1951).

the linguistic code, in general, that to a large extent determines our perception of the speech sounds. We perceive them not as mere sounds but specifically as speech components. More than this, the way we perceive them is determined by the phonemic pattern most familiar to us.' That is, a phonemic (or any linguistic) pattern is precisely the sort of 'shared frame of reference' that Newcomb speaks of. And to go one step further: just as Hjelmslev speaks of the process presupposing the system, so Newcomb asks for the study of 'those kinds of interaction which presuppose shared frames of reference, and by which shared frames of reference are acquired. This special kind of interaction I shall refer to as communication.'

2. THE STATUS OF MEANING IN LINGUISTICS. Jakobson-Fant-Halle use semantic criteria to define their terms (14): 'The smallest meaningful unit in a language is called *morpheme* ... The distinctive feature and the phoneme possess no meaning of their own. Their only semantic load is to signalize that a morpheme which, *ceteris paribus*, exhibits an opposite feature is a different morpheme.' Phonemes and distinctive features are thus, to them, not simply recurrent partials and their components,¹² but units having a definite functional load; the entire presentation of the *Preliminaries* is in terms of the discriminative function of the phoneme and the distinctive feature, which is quite clearly based upon a semantic assumption. This follows logically from the initial assumption treated above, namely that language is an existing system—a system of signs. The sign, in Saussurian terms (op.cit. 99), is a joining of the signifier and the signified, that is of form and meaning, into an indissociable unit; hence, if the linguistic description is to picture the reality of language correctly, it has to proceed, not merely in terms of both form and meaning, but primarily in terms of the interrelationship between form and meaning. This is the core of de Saussure's argument; it is also the rationale, for instance, for Jakobson-Fant-Halle's basic technique of isolating phonemes: the mutual substitution of phonemic segments in identical environments to see whether such a substitution (commutation) entails a differentiation of meaning (1-2). The commutation differs very little from any other standard technique of identifying phonemes; the difference lies in the overt use of meaning as a criterion of identification.

This is, of course, the most controversial aspect of Jakobsonian theory.¹³ Still, in using substitution techniques for identifying linguistic units, linguists of all shades of opinion rely on the criterion of sameness vs. difference. Thus, Bloomfield states in the 5th definition of his Postulates of 1926:¹⁴ 'That which is alike will be called *same*. That which is not same is *different*.' This is rephrased by Bloch (op.cit. 7) to read: 'Some utterances contain auditory fractions that are the same.' The crucial point, it seems to me, is the kind of sameness. Bloch

¹² For the contrary approach see Harris, *Lg.* 20.185-8. Here any mention of distinctive or other semantic function is avoided.

¹³ Cf. Bloch, *Lg.* 24.5: 'In our wording we shall avoid all semantic and psychological criteria. The implication is, of course, that such criteria play no part, or at least need not play one, in the theoretical foundation of phonemics.'

¹⁴ Leonard Bloomfield, *A set of postulates for the science of language* (1926), reprinted *IJAL* 15.196 (1949).

is willing (op.cit. 8) to 'grant that this sameness is no doubt ultimately a matter of biosocial equivalence', but would 'prefer to account for it simply in terms of recurrent auditory fractions that are (by assumption) the same.' Nida is less reluctant to define sameness in semantic terms when he speaks of the 'common semantic distinctiveness' of morphemes.¹⁵

The inclusion of meaning in a definition of the phoneme (and an operational procedure for isolating phonemes) does not presuppose a semantic analysis; it merely requires the admission that sameness in language is a matter of functional equivalence rather than physical identity. Both older and more recent experimental phonetic procedures have served to point up the fact that even 'same phonations'¹⁶ are not identical in either frequency count or articulatory characteristics. Sameness can thus, for linguistic purposes, be said to exist in terms of perception, which is to say—as I have pointed out above—in terms of the social norms determining the perception. To go one step further: how shall we know whether two phonetically similar segments in identical environments are linguistically 'same' (i.e. free variants of the same phoneme) or 'different' (i.e. different phonemes)? Phonetic criteria will not be very useful here, and distributional criteria will lead to an indefinite extension of the environment¹⁷ which is operationally undesirable since it contradicts the requirement of simplicity.¹⁸ It seems to me that much complicated theoretical reasoning can be avoided by accepting the fact that linguistic sameness fundamentally boils down to some such criterion as Nida's common semantic distinctiveness. Jakobson-Fant-Halle assert no more than this, since they use meaning difference, rather than a statement of meaning itself, as their differentiative criterion.

3. THE STATUS OF PHONETIC DATA IN PHONEMICS. The bulk of the *Preliminaries* is built around a positive conception of the status of phonetic data in phonemics. This point of view is much more controversial among European linguists than in America. The most clearly contrary assumption to that of the authors is Hjelmslev's; he considers the phonetic (or any other) substance irrelevant to linguistic form, which is the only legitimate subject matter of linguistic analysis: 'We must be able to imagine as ordered to one and the same linguistic form substances which ... are essentially different; the arbitrary relation between linguistic form and purport makes this a logical necessity' (op.cit. 66). This has stirred up quite a hornet's nest of discussion among colleagues in Europe.¹⁹

Quite a few American linguists have likewise been reluctant to use phonetic data for phonemic classification, and have quite consistently preferred a treatment in terms of distribution. Thus, Harris calls his analysis of simultaneous components (op.cit. 181 fn. 6) 'primarily distributional rather than phonetic'.

¹⁵ Eugene A. Nida, *Morphology* 7, 14, and passim (Ann Arbor, Mich., 1949).

¹⁶ Bloch, *Lg.* 24.10.

¹⁷ Harris, *Methods in structural linguistics* 156 ff. (Chicago, 1951).

¹⁸ Hjelmslev, op.cit. 6.

¹⁹ Cf. Josef Vachek, Some remarks on writing and phonemic transcription, *Acta linguistica* 5.86-93 (1945/9); Fritz Hintze, Zum Verhältnis der sprachlichen Form zur Substanz, *Studia linguistica* 3.86-105 (1949); and the reviews of Hjelmslev's *Omkring sprogteoriens grundlæggelse* by Eli Fischer-Jørgensen in *Nordisk tidsskrift for tale og stemme* 7.81-96 (1944) and by André Martinet in *BSL* 42.19-42 (1946).

Phonemic patterns are by most American authors described in precisely those terms, but phonetic data—as opposed to considerations of meaning—are accepted as part of the description.²⁰

How essential phonetic data are to Jakobson's phonemic argument appears most clearly from his interpretation of children's language and aphasia,²¹ where he presents phonemic structure as a hierarchic order conditioned by the properties of the phonetic substance. Several students of children's language have confirmed the general outlines of Jakobson's systematization, among them such a thorough worker as W. F. Leopold.²²

To most American linguists phonetic means articulatory. The use of acoustic data, as Hockett puts it (*IJAL* 13.259), 'might be highly desirable if it were possible, but as yet, despite a number of efforts in that direction, it seems not realistically feasible.' In Jakobsonian terms, an acoustic approach is not merely highly desirable, but essential to a proper interpretation of phonemic patterns. This, I believe, is primarily due to Jakobson's conception of langue as a perceptual system. Jakobson-Fant-Halle state this quite clearly (12): 'In decoding the message received (A), the listener operates with the perceptual data (B) which are obtained from the ear responses (C) to the acoustical stimuli (D) produced by the articulatory organs of the speaker (E). The closer we are in our investigation to the destination of the message (i.e. its perception by the receiver), the more accurately can we gauge the information conveyed by its sound shape. This determines the operational hierarchy of levels of decreasing pertinence: perceptual, aural, acoustical and articulatory (the latter carrying no direct information to the receiver).'

In addition to this epistemological justification, the use of acoustic rather than articulatory criteria allows the authors to find additional symmetries in the patterning of phonemes. I refer here primarily to the use of compactness and tonality in both consonants and vowels: *k* contrasts with *p* as compact versus diffuse, just as *a* does with *ə* on the basis of the same feature; likewise grave *p* contrasts with acute *t*, just as grave *u* does with acute *i* (27-30). These parallelisms do not appear when consonants are classified in the customary manner by points of articulation; this was pointed out by Jakobson in a much earlier paper,²³ but the *Preliminaries* go considerably further in establishing detailed distinctive features than any previous publication.

The acoustic orientation of the authors constitutes a rather important departure from Trubetzkoy's *Grundzüge der Phonologie*,²⁴ which is usually considered the major source of information on Prague School phonemics. Trubetzkoy classifies phonemic contrasts primarily on the basis of articulatory criteria, and

²⁰ Cf. Bloch's formulation, loc.cit.: 'Theoretically it would be possible to arrive at the phonemic system of a dialect entirely on the basis of phonetics and distribution, without any appeal to meaning . . .'

²¹ Jakobson, *Kindersprache, Aphasie und allgemeine Lautgesetze* [1941] (Uppsala, 1942).

²² *Speech development of a bilingual child* 2.197 ff. (Evanston, Ill., 1947).

²³ Observations sur le classement phonologique des consonnes, *Proc. 3d Internat. Congr. of Phon. Sciences* 34-41 (Ghent, 1939).

²⁴ *TCLP*, Vol. 7 (Prague, 1939). Page references are to the French translation by J. Cantineau, *Principes de phonologie* (Paris, 1949).

arrives at a rather cumbersome system (69 ff., esp. 135 ff.). One aspect of this system, in particular, is considerably simplified by the procedure of the *Preliminaries*: in accord with the relationship between the contrasting members, Trubetzkoy divides phonemic oppositions into privative (presence vs. absence of a characteristic), gradual (different degrees of a characteristic), and equipollent (one characteristic opposed to a different characteristic, such as the different points of articulation, 77), and considers the latter in any system the most frequent of all. From the standpoint of systematization, however, the privative oppositions are the most fruitful, since they lead to the establishment of correlations (87-93), and are at the basis of the Trubetzkoyan treatment of neutralization (87).

Jakobson-Fant-Halle's treatment in acoustic terms eliminates the 'equipollent' category which Trubetzkoy had used, for instance, in interpreting phonemic contrasts based on different points of articulation. These now become 'privative' contrasts based on compactness and tonality (27-30). Trubetzkoy's 'gradual' oppositions are also assimilated to the 'privative' by replacing several degrees of a given characteristic by contrasts based on two features, each of these being privative (e.g. compact vs. non-compact, diffuse vs. non-diffuse; 29). A trace of the distinction between 'gradual' and 'privative' oppositions remains in the distinction 'either between two polar qualities of the same category ... or between the presence and absence of a certain quality' (3).

In essence, all phonemic oppositions become 'privative' in the Trubetzkoyan sense, or at least near-privative, and can thus be included in both correlations and neutralization. This not only simplifies the Trubetzkoyan classification, but increases the scope of some of its most useful concepts, controversial though they may be.

4. THE STATUS OF PHONEMIC CONTRASTS. In my discussion of the preceding part, I have already touched upon Jakobson's tendency to present linguistic patterns in terms of binary contrasts. But the technique used by the authors to systematize oppositions into this binary pattern is here, like their technique in general, primarily the implementation of an epistemological insight. Phonemic patterns are interpreted in terms of binary contrasts, not for the convenience of the linguist, but because (9) 'in the ... case of speech ... a set of binary selections is inherent in the communications process itself ... The dichotomous scale is the pivotal principle of the linguistic structure. The code imposes it upon the sound.' The binary interpretation of phonemic patterns thus take on the nature of an epistemological postulate, and the emphasis on acoustic data allowing most clearly for such an interpretation becomes a corollary to this postulate.

It is worth noting that Jakobson's interpretation of phonemic patterns in binary terms was launched well ahead of a detailed acoustic description. In his paper of 1939 (fn. 23) the emphasis is on a reinterpretation of articulatory data in terms of resonator shapes and orifices, rather than the customary points of articulation; in his monograph of 1941 (fn. 21), the dichotomies apparent from the learning process of children, and the decay of the language of certain aphasiacs, are presented in synesthetic terms. This is, of course, perfectly good

scientific procedure: on the basis of certain data (mostly those on children's language and aphasia) to formulate a hypothesis (that of the dichotomous structure of phonemic patterns), and to try to find corroborative evidence by introducing new sets of data: reevaluated articulatory facts, synesthesia, acoustic information.

The evaluation of the binary theory then boils down to two basic questions: (a) How reasonable is the hypothesis? (b) How conclusive is the support given to it by the corroborative evidence?

Concerning the reasonableness of their basic assumption the authors provide part of the answer by linking the binary hypothesis to information theory (9), which 'uses a sequence of binary selections as the most reasonable basis for the analysis of the various communication processes.' Information theory thus provides the most important theoretical parallel to the binary hypothesis so far discovered. In this sense the binary hypothesis is accepted as at least a reasonable alternative by some workers on the 'other side of the fence', among them Licklider:²⁵ 'Essentially, the binary distinctions provide a way of simplifying the patterns on which the identifying mechanism [of communication] acts. Since the distinctions are to some extent variable and since they contain less information than the more complete patterns from which they were derived, one would expect an identification process based on binary distinctions to make considerable use of conditional probability structures in deciding among alternatives not clearly eliminated by information carried by the current set of distinction.' Licklider also neatly summarizes the status of Jakobson-Fant-Halle's corroborative acoustic evidence (op.cit. 590): 'It is not yet entirely settled what the acoustic counterparts of some of the phonatory distinctions are, but it is nevertheless of interest to picture speech as the product of an encoding process that is governed by a set of decisions between paired alternatives.' This is an evaluation which I would be tempted to accept. As the title of the *Preliminaries* suggests, the authors do not claim to present a definitive statement of their hypothesis. Some of the interpretations are obviously better grounded than others, and are presented with correspondingly greater emphasis.

There are, of course, not a few phonemic situations in which binary contrasts do not readily appear; this is particularly true of some prosodic features. One case with which I am somewhat familiar involves the five levels of pitch in Trique, a Mexican Indian language described by Longacre.²⁶ When I discussed the matter with him in 1949, it appeared that the situation was very simple, and clearly binary: four pitches on two levels, with high-low pitch for each level. Since then, Longacre has brilliantly demonstrated a fifth pitch level in the language, but his summation of the pattern still leaves the way open for a binary interpretation (op.cit. 80-1): 'The tonemic oppositions may be conceived of as consisting essentially in a bi-directional spread from a relative norm. This centrifugal opposition is actualized in five relative levels. The center level, lying

²⁵ J. C. R. Licklider, On the process of speech perception, *Jour. Acoust. Soc. Amer.* 24.594 (1952).

²⁶ Robert E. Longacre, Five phonemic pitch levels in Trique, *Acta linguistica* 7.62-82 (1952).

quite close to the phonetic norm, may be considered the structural norm of the system. The highest level represents one pole of the bi-directional tendency; the lowest level represents the other pole. Between norm and high occurs a further level raised from norm; between norm and low occurs another level lowered from norm.' In Jakobsonian terms, the Trique tonal system thus has two distinctive features (high and low) distributed over the five tones as follows:

	TONE 1	TONE 2	TONE 3	TONE 4	TONE 5
HIGH	+	±	-		
LOW			-	±	+

I now turn to some aspects of the *Preliminaries* that are not directly related to the four questions posed above.

As I mentioned earlier, Jakobson-Fant-Halle not only do not neglect distribution in favor of other phases of phonemic interpretation, but use distributional data as an essential part of their argument, namely in their discussion of Invariance and Redundant Variations (4-8). Here, quite obviously, distribution is used to distinguish between phonemic invariance and allophonic redundancy. Their discussion is closely intertwined with the approach in terms of distinctive features, especially in the interpretation of the fortis-lenis contrast in Danish and English: it is not, to them, a matter of the distribution of individual phonemes, but of phonemic oppositions. They consider, for instance, the contextual representatives of both members of an opposing pair considered together; in Danish, this leads to the assignment of phonetically similar items to opposite ends of the contrast in differing contexts: phonetic [d] represents phonemic /d/ in the 'strong' position, phonemic /t/ in the 'weak' position.

The matter of invariance and redundancy is not limited to phoneme distribution in the linear sequence, however. 'Another less analyzed though very important class of redundancies is conditioned by the superposition of simultaneous distinctive features. There are languages in which the velar [k] is in complementary distribution with the corresponding palatal stop ... By the same reasoning, if in French we find the velar stop /k/, the palatal nasal /ɲ/ (as in *ligne*) and the prepalatal constrictive /ʃ/ (as in *chauffeur*), we must consider the difference between this velar, palatal and prepalatal articulation as entirely redundant, for this difference is supplementary to the other, autonomous distinctions' (6).

Again, in the charting of distinctive features in the appendix on Analytic Transcription, the distribution of features within the system is used to eliminate redundancies and simplify the charted pattern (43-5). Here, as in other aspects of their work, the authors continue in the Trubetzkoyan tradition: the *Grundzüge* (39 and passim) quite clearly differentiate between phonemically pertinent and phonemically 'non-pertinent', i.e. redundant, features.

The authors' choice of some of their specific data will undoubtedly be criticized by specialists. Many linguists interested in English will be unhappy over the use of 'Received Pronunciation'. Jakobson promises to discuss his analysis of English in more detail in his forthcoming book *Sound and meaning* (vi); his reasons for choosing 'Received Pronunciation', and for the particular interpretation of it in the *Preliminaries* (as hinted at, 43-4), will then perhaps be more

apparent to his critics. Although Jakobson-Fant-Halle strive to eliminate the non-uniqueness of phonemic solutions by the distinctive feature technique (7), there will probably always be room for disagreement as to the analysis of specific data.

It is equally clear that the theoretical foundations of the *Preliminaries* will remain controversial for quite some time. It is my feeling that the controversy is a healthy one, for it injects some useful leavening into the dough of theoretical discussion. It is, incidentally, a discussion which has lately grown far beyond the disciplinary confines of linguistics, and has led to many new fruitful insights and points of departure.

It will be up to linguists more familiar with acoustic and experimental phonetics to evaluate the details of the authors' argument, but I would venture to say that the authors, however tentatively, have proposed a useful organizing principle to interpret spectrographic and other acoustic data. The validity of the principle rests, in my opinion, not on the correctness of any one detail of interpretation, but on the soundness and logical acceptability of Jakobson's basic assumptions and techniques. These are what I have attempted to discuss.

Indogermanisches etymologisches Wörterbuch. By JULIUS POKORNY. Fascicle 7, *ker-* to *leizd-*, pp. 577-672. Bern: A. Francke AG, Verlag, 1953.

Reviewed by JOSHUA WHATMOUGH, *Harvard University*

I have no embarrassment that hinders me in reading a dictionary, and it is a reviewer's duty to read a book that he reviews. But readers of the review are more likely to consult the dictionary at need. As I have said before, each will find something to learn from Pokorny's *Idg. etym. Wb.*, to add to it, or to correct. The following remarks are a running commentary of one reader (or user) only, though I am in private honor bound to say that I have tried to be catholic, within the restrictions of one man's competence, in my comments.

In his preface Pokorny assures us that everything of importance that appeared before the end of 1947, so far as available to him, has been considered. Actually later references are given to more recent publications, e.g. to J. B. Hofmann's *Etymologisches Wörterbuch des Griechischen*, which appeared in 1949. But Pokorny's work preserves its old-fashioned flavor. For 'heart' (579) we have (*kered* :) *kērd*, *krd*, *kred*, precisely as in Walde-Pokorny 1.423 (except that it has *kērd* instead of *kērd*, which may well be accidental). Since Pokorny will have none of the newer theory, he will not write **keHrd-*, as most of us now do, with Greek *κῆρ* and the Baltic accentuation in mind. As for *crēdō* (580), the attempt to trace its semantic history made by Dumézil, *Revue de philologie* 64.313-7 (1938) (comparing Armenian *aril'*), is entirely passed over. But so it was by Ernout (1951), whose *Dictionnaire étymologique de la langue latine* calls itself 'histoire des mots', and who gives a long paragraph slavishly copied from *MSL* 22.218 (1921). On the same item Pokorny refers to Vendryes, *RC* 40.436 (1923); Ernout to *RC* 44.90 (1927). This is a mystery too deep for me; unless Dumézil is deliberately despised by Pokorny and by Ernout, as well as by others.

The local name *Cremona* is connected with Greek *κρόμμον*, commonly translated

'onion', though there is something to be said for 'leek' or even 'garlic'; this is possible, for Stokes connected *Cularo* (the modern Grenoble), plausibly enough, with Ir. *cularan* 'cucumber' (denied by Walde-Pokorny 1.354), like Greek Σελών, cf. σίκκος 'cucumber'.

The Greek ξέστρς (585) 'hewn, planed' is the word traditionally read (ξέστρων) at Mark 7.4 (*ligre* in Ulfilas, *beds* in most English versions); actually the ξέστρων of the New Testament is a different word ξέστρς (-ου) 'pitcher, cup', the etymology of which, according to Liddell and Scott, is nothing more than Latin *sextārius*. But it is not easy to see how *sextārius*, borrowed into Greek, would give ξέστρς, and a wooden cup is manufactured precisely by gouging (ξέω), and Greek not only had, at least as early as the 6th century, both ἐκρεῖς and ἐκρή for a liquid measure *sextārius*, but regularly wrote σ- for Latin s- (e.g. σεκουνδα-pοδῆς from *secundā rude*, σεκοῦπιον from *secūris*). There are other difficulties in the assumption that ξέστρς is a Hellenized *sextārius* (see Schwyzler, *Gr. Gr.* 1.299). The borrowing of Latin (or Italic) words into Greek, in its chronological as well as phonological and morphological aspects, needs further study.

It is a misfortune that Pokorny, though he regularly gives references to Walde-Pokorny and to Walde-Hofmann, gives none to Ernout-Meillet's second edition, if not the third. Under *kes-* 'schneiden', *k-s-tro-m* 'Schneidewerkzeug, Messer' (586) we have an interesting series of Italic words: Latin *carēre*, *castus* (but *castitās* is omitted), *castigāre*, *cassus*, *castrāre*, *castrum*, Oscan *kasit*, *castrous* (gen. sg. 'fundī'), Umbrian *castruo*, *kastruwuf* 'fundos'. There was an opportunity here to give a hint of the semantic string, or at least a reference to a place where a discussion of it (other than Walde-Hofmann 1.167, 178 ff.) might be found. See now Ernout-Meillet's 3d edition 184-6 (1951), where the possibility is considered that *recesse* also is involved. The ancients not unnaturally derived *castus* from *castrāre*, but that of course is wide of the mark. The castrated are by no means chaste, notwithstanding Origen's goings on. And Housman was right when he said that 'of all sexual aberrations chastity', i.e. total abstinence, 'is the rarest.'

The α- of ἀκούω (587) is a nut that still refuses to be cracked. The best guess is that it will turn up as *H* one of these days; cf. Hammerich, *Laryngeal before sonant* 29 (1948).

589: This etymology of *cubus*, κύβος is correct. How far are historians of mathematics correct in seeing the influence of Arabic *ka'b* in *cube* (root etc.)? Lokotsch agrees in ignoring that hypothesis. Or did Greek influence Arabic in this case? Kretschmer, *KZ* 55.89 (1928), thought he could trace a Lydian word in *cubus* 'gaming die'.

597: The alleged connexion of Greek κυκεών and Lithuanian *šdukštas* is indeed 'sehr unsicher', but it escapes the forbidden pattern *kek-* which appeared on p. 543. As Benveniste showed, in IE *CvC* the consonant may be of any order, provided that the two orders differ, e.g. *kel-* and *lek-* occur, but not *kek-* or *lel-*.

612: Pokorny is one of those who maintain that Ligurian was not Indo-European. Yet here we have (s.v. *kom* 'mit') *ligur*. FIN. *Comberanea*. According to Pokorny's own theories (which I do not accept), this should be labeled 'illyr.', not 'ligur.' Similarly (623) Gallic *cruppellarii* (with which *Mons Graupius*, the

modern *Grampians*, has been connected by Pokorny, reading **Croupius*) is 'wegen des *p* veneto-illyr. Herkunft'. Whatever the explanation of the *p* or *pp*, it is a long, long way from Illyricum to the Highlands of Scotland. The word *cruppellarius* is known from Tacitus *Ann.* 3.43; it is applied to armored gladiators; and a better case could be made for a Germanic source (cf. *grob* 'uncouth'). More than one German word (e.g. *framea*) got into ancient Gaul.

625. For ξηρός 'dried fruits' cf. my review of Onions' *Origins of European thought*, forthcoming in *Classical philology*. As for *k̑p-* (two items only, 626), Buck's remarks in *Lg.* 2.102 (1926) are pertinent. 626: There are better etymologies of *kiss* than the assumption 'Schallwort'. Cf. rather *gustāre* 'taste'. 631: For *combrētum* : *combri* see now *DAG* 178; and the same place for *prenne* 'arborescens' (633). The Endlicher glosses, as Thurneysen showed, *IF* 42.143-6 (1924), contain local names; cf. therefore *Compriniacus*. The Coligny *prinni* is possibly cognate. 648: Add Keltic *ruv(v)* 'for whom', *DAG* Note lviii. 652: Add perhaps Venetic *leivovs*, *PID* 152. 653: A very large number of examples of *o* : *a* (especially after resonants) has turned up in *DAG*, in addition to the well-known *locus* 'lake', *more* 'sea', with which I hope to deal elsewhere. 653: On *lama* cf. *lamaticom* in Spain (*CIL* 2.416; *DAG* *247, 1346). 654: *lat-* beside *let-* (W-P 2.427) would appear to be a case of distinctive *H*₁ and *H*₂. But is it possible that the three (or two?) roots *leg-* had distinctive *H*₁ and *H*₂ and also *H*₃ initially (before *l-*)? Pokorny need not face this question, but those who believe in laryngeals cannot escape it.

Westtocharische Grammatik 1. Das Verbum. By WOLFGANG KRAUSE. (Indo-germanische Bibliothek; 1. Reihe: Lehr- und Handbücher.) Pp. xvi, 312. Heidelberg: Carl Winter, Universitätsverlag, 1952.

Reviewed by GEORGE S. LANE, *University of North Carolina*

This indispensable volume is the first of a contemplated complete grammar of 'Westtocharisch', as the author has chosen to call that dialect which has been traditionally known in German as 'Tocharisch B', in England as 'Kuchean', and correspondingly in French as 'Koutchéen'. In conformity thereto, 'Tocharisch A' is here designated as 'Osttocharisch'. The names are good ones, but have the misfortune of being unfamiliar, and, since the area of their provenance is, I should presume, even more unfamiliar to most scholars than the languages, they will no doubt cause some confusion until we get used to the fact that Kucha is west of Turfan. Personally of course I prefer to keep 'Kuchean', which is familiar in England and France and not unknown in German, and to adopt for Tocharian A a parallel 'Turfanian', from the name of the place where the documents are chiefly found.¹ It is a great misfortune for the layman that we who are concerned with this field have not been able to agree on our terminology even after forty odd years of research.

This volume consists of five principal and quite unequal parts. The first is concerned with matters of writing and phonology (A. Zur Schreibung und Laut-

¹ I regret that I did not do this in my recent *Studies in Kuchean grammar I* (Supplement to *JAOS*, No. 13). There I used 'Tocharian' simply for Tocharian A.

lehre: I. Zum Vocalismus; II. Zum Konsonantismus, 1-22). The second part (B. Der Aufbau des westtocharischen Verbalsystems, 23-58) gives a broad account of the structure (voices, moods, tenses, derivative systems, non-finite forms, etc.) of the Tocharian verb. The third part (C. Bildung der einzelnen Tempora und Modi, 59-192), is the main part of the book. Here are outlined the various classes of present (and imperfect) and of subjunctive (and optative), the imperative, the durative, and the preterit. The fourth part (D. Die Personalendungen, 193-203), besides listing the endings by classes, makes a brief attempt to account for them historically. Finally the fifth part (E. Vollständige Paradigmata, 208-16), gives complete representative paradigms of the tense classes and modes. The rest of the volume (217-309) is a verbal index, comparable to that at the end of SSS.²

The first part rightly makes no attempt to give a detailed 'Lautlehre' from a historical or descriptive point of view. Instead, it deals with problems of orthography or phonology or both, which are certain to confuse the beginning student of Kuchean, even if (or perhaps, indeed, partly because) he has had some first acquaintance with Turfanian, where the phonology, at least as it is reflected in the orthography, is much more regular. In Kuchean, the orthography is subject to variation not only in one and the same manuscript, but especially in manuscripts of different provenance. And among the manuscripts of the Berlin collection, those which come from the general area of Kucha, especially from Ming-öi Qizil (MQ) and from Qumtura (Qu), show the widest deviations from those found at Turfan and Qarašahr, i.e. in the 'Turfanian' area. We are therefore indebted to the author for giving us a survey of these apparently purely orthographical stumbling blocks at the very outset.

In this section the author treats also the *o* movable (bewegliches *o*) which may be found, according to the demands of meter in poetical texts, after any final consonant in all categories of forms, except that it is not found (9 bottom, ftn. 2) after the ending *-m* in the 3 sg. or pl. of verbs 'weil diese Endung ersichtlich auf die indogerm. Secundärendung *-nt* (also ohne schliessenden Vokal) zurückgeht ...' Would the author imply by this statement that the *o* has elsewhere an etymological value, i.e. reflects (or replaces) a final IE vowel of any sort? If so I am very doubtful.

Here too is a treatment of word accent and the attendant theories concerning vowel weakening. In fact Krause believes it possible to derive the rules of accent for Kuchean from the apparent vowel weakenings in unaccented syllables. The chief rule is that the accent fell on the penult in words of two or three syllables. This would be indicated by vowel weakening (*a* > *ā*, *ā* > *a*) before and after the accent: *kante* '100' : pl. *kāntenma*, *tāka* 'was' : *takāre* 'were', and the like. For words of more than three syllables, however, no rules are apparent; and even for words of three syllables there are so many exceptions that I feel sure (as I believe the author does also) that the final word has not yet been said about Tocharian accent.

The first section of the grammar contains also a valuable survey of palataliza-

² SSS as usual stands for Sieg, Siegling, and Schulze, *Tocharische Grammatik* (Göttingen, 1931).

tion (Erweichung). The author holds to his previous view³ that the phenomena must be separated into assibilation ($t > ts$), which occurred under unknown conditions in pre-Tocharian times, and actual palatalization ($t > c$, $k > ś$, etc.), which took place originally before palatal (front) vowels. In addition to the common Tocharian changes, Kuchean shows further a secondary analogical palatalization of $k > ky$, $p > py$, $m > my$, $ts > tsy$ in the preterit of the causative. Only now, with a complete catalog of verbal forms in our hands, can we see the function which palatalization plays in the Kuchean conjugation.

A few of the matters considered in the second part seem to call for some comment. The author stresses, as is to be expected, the contrast between the primary verb (Grundverb) and the causative. The value of the *sk* (§§) formation as a causative is of course not to be denied, but on the other hand it must not be overemphasized, i.e. it hardly seems justifiable to make the statement (24) that 'bei einigen auf *k* endenden Wurzeln sind nur kausative Bildungen vorhanden'. The verbs cited in this connection—*nakšām* 'destroys', mid. *nakštār* 'perishes'; *pakšām* 'ripens (trans.), cooks', mid. *pakštrā* 'becomes ripe, is cooking'; *tsakšām* 'burns (trans.)', mid. *tsakštrā* 'is burning (intrans.)', etc.—show, in my opinion, not a weakening or loss of a fully developed causative meaning, but rather the point of its inception. The meaning of the *sko* suffix, while not to be fixed absolutely, is clearly to be sought in the iterative (intransitive) and factitive (transitive) functions in many Indo-European languages. It would seem to me plausible that the causative use of the suffix developed out of the intransitive middle use in just such verbs as are cited above, e.g. intransitive 'perishes', hence active 'destroys' = 'sends to perdition, causes to perish'. In such cases as *tāllām* 'lifts' beside *talāššām*, which is according to Krause 'eine formale Kausativbildung ohne erkennbar kausative Bedeutung', it would seem better to depart from the more archaic values of the *sko* suffix rather than from the late Tocharian use. In general, though much detailed work must still be done, it seems clear to me that Kuchean is much more archaic than Turfanian in the use of this formation. That is, the development of a uniform causative value for the *sko* formation (Kuch. *sk/§§*, Turf. *s/§*) is much further advanced in the latter dialect.

This same section contains also (26–41, §§19–32) a brief syntax of the verb, including both the finite and the nonfinite forms. This is indeed most welcome, and will serve as a stopgap until the necessary researches have been done to write a real syntax of the two dialects. Even from this brief survey of the function of the tenses, however, I am further confirmed in my suspicion that we have been much too hasty in our judgment of the difference of values between what has been called the imperfect and the preterit tenses in Turfanian. In a detailed study in the Edgerton number of *LANGUAGE* I have made it clear, I believe, that the Turfanian imperfects are identical in origin with the preterits of Turfanian itself or of Kuchean or of both. I have also intimated that the difference in function (i.e. in grammatical meaning) has in all probability been exaggerated. The situation in Kuchean would, I believe, bear out that conclusion. Here the Proto-Tocharian 'imperfect' of optative origin, which is formally quite distinct from the

³ *Gött. gelehrte Anzeigen* 1943.30 ff.

preterit, is clearly on the wane. In Turfanian it is lost (except in the case of the two verbs 'be' and 'go'). A random glance at the verbal index of the grammar or at the listing of the imperfect forms (104-9) over against that of the preterit (159-92) clearly reveals the disproportionate use of the latter. In Turfanian, in fact, considerably less than a hundred different imperfect forms, not counting the verbs 'be' and 'go', are listed in SSS.

The syntax of non-finite forms (33-55) deals not only with the use of the infinitive, participles, gerundives, and absolutives, but also with almost every sort of verbal derivative, such as privatives (e.g. *akākatte* 'uninvited': *kāk-* pret. root to *kwā-* 'call'), nomina agentis, verbal adjectives, adverbs, and substantives of all sorts if derived from the verbal root or from other verbal forms. In addition there is included a most welcome discussion of periphrastic formations, both with the preterit participle (e.g. *sanune kekamu nesau* 'I have come into danger') and with the gerundives (e.g. with gerundive I *šaumontse mā ste waike weşşalle* 'one must not tell a lie', lit. 'of men a lie is not to be told'; with gerundive II *mā po yātalle sāmñe kantwasa weṃtsi* 'it is not possible to tell all with human tongue'). In his appraisal of the difference in value of the two gerundives (I from the present stem, II from the subjunctive), Krause follows the thesis of Werner Thomas in his monograph, *Die tocharischen Verbaladjective auf -l*.⁴ he attributes the notion of necessity or imperative value to I, of mere possibility to II (cf. the illustrations above). Gerundive II is also used with the present copula in the value of a future, especially with a negative; e.g. *mā tañ kc-āyor aille nesau mā-lyekepi* 'I will not give you or anyone else a gift'. With the imperfect of the copula we get an unreal condition: *akālkā tsānkā-ne makte pi kca tā onkorñai ñiś šwātsi källalle şeym* 'there arose in him the desire: how might I be able to get this porridge?'

The third part of the grammar, as I have already indicated, is the meat of the book. It is only here that one begins to realize, or at least has again forcibly recalled to him, how different the two dialects really are. This difference extends throughout the verbal system. Even the present-stem formations can be only approximately compared. It would be impossible to set them up in an approximately parallel fashion, as is the case for example in such apparently widely divergent languages as the Germanic languages, or as Oscan, Umbrian, and Latin.⁵ Only in a few instances are there comparisons that are of significance to the classification as such. A brief outline is in point.

Kuch. Cl. I (= Turf. Cl. I, SSS 350 f.), apparently originally the athematic class, but with adopted thematic vowel *ā/e* in the active but without palatalization before *ā* (as in thematic verbs, cf. Cl. II below), e.g. sg. 3 *palkām* 'shines' (= Turf. *pālkāş*), pl. 3 *salpeṃ* 'glow' (= Turf. *sālpiñc*). In the middle, so far as the evidence goes, the 'thematic' vowel *a* or *ā* is adopted in the 3 plural (*kolokan-trā* 'they follow', *pkwāntār* 'they trust'), in the verbal adjective (*pkwalle* to *pākw-* 'trust', *tsopalle* to *tsop-* 'prick'), and optionally in the middle participle (*kolokmane*, *sālpamane*).

Kuch. Cl. II (= Turf. Cl. IV, SSS 353 ff.), the thematic class, with inherited

⁴ Berlin, 1952.

⁵ Cf. Buck, *A grammar of Oscan and Umbrian* 9 ff., 151 ff.

thematic vowel *ā/e* < PIE *e/o* and palatalization (if possible) before *ā*, e.g. from *āk-* 'lead': act. sg. 3 *āśām*, pl. 3 *ākem*, pple. act. *āseñca*, mid. *ākemane* (= Turf. act. pl. 3 *ākeñc*, mid. sg. 3 *āstrā*, pple. act. *āsant*); from *aik-* 'know': mid. sg. 1 *aikemar*, 3 *aistrā*, pl. 3 *aikentrā*, pple. act. *aiseñca*, mid. *aikemane*, vbl. adj. *aisalle*. In this class are included of course thematic formations other than the simple *e/o*-suffix type, e.g. act. pl. 3 *keriyem* 'laugh', pple. mid. *keriyemane* (= Turf. *kāry-* Cl. III), clearly belongs to the *ye/yo* class (< PIE **ghr̥ryé/yo*, cf. Gk. *χαίρω*), and so also probably does mid. sg. 3 *kaltār* 'stands', pl. 3 *klyentār* (= Turf. *kāly-* Cl. IV), perhaps from PIE **k^wélye/yo* (cf. Gk. *περι-τέλλομαι* 'come around'), though the absence of palatalization of the initial (*k* > *ś*) is strange. Other verbs listed here are original *sko*-verbs, e.g. *pāsk-* 'keep' (cf. Lat. *pāsko*, *pāvī*, etc.), *nāsk-* 'bathe', *trāsk-* 'chew' (cf. Lith. *dreskiù*, *drėskiaũ*, *drėksti* 'tear', iter. *draskaũ*, *draskyti*?—Walde-Pokorny 1.8-3).

It is of course frequently impossible to tell whether or not a particular verb belongs to Cl. I or II, either because the key forms which would show palatalization (sg. 2, 3, pl. 2, and act. pple.) are not attested or because the final consonant is 'unpalatalizable'; e.g. pl. 3 *trusen-me* (*trus-* 'lacerate') in the former category and sg. 3 *campām* etc. (*cāmp-* 'be able') in the latter.

Kuch. Cl. III shows a stem vowel *e*. The majority of cognate verbs in Turfanian are listed in SSS as Cl. II presents (*a* stems), but a number are Cl. III (= Kuch. Cl. IV, *o*-verbs, cf. below). A good example is *māsk-* 'be': sg. 1 *māskemar*, 3 *māsketār*, pl. 3 *māskentār*; pple. act. *māskeñca*, mid. *māskemane*; vbl. adj. *māskelye*. Except for *māskeñca*, all quotable forms belonging to this class, finite or nonfinite, show middle endings only. Turfanian Cl. II verbs have both active and middle forms. Historically these verbs would seem to show already in Proto-Tocharian an extension of the *o*-grade thematic vowel to all persons. This extension must have occurred prior to the palatalization, since there is no trace of palatalization in them.

Kuch. Cl. IV contains *o*-verbs. Besides showing the characteristic *o* where Cl. III has *e*, this class mutates a radical *a* or *ā* to *o*. Radical *ai* and *au*, the only other two vocalisms found, are not mutated. The majority of Turfanian cognates belong to SSS Cl. III, which, like Cl. II (= Kuch. Cl. III), shows the stem vowel *a*, but which, unlike II, drops it before the endings *-mār*, *-tār*, *-mām* (cf. SSS 352 f.). Examples are Kuch. *ār-* 'cease', mid. sg. 3 *orotār* (Turf. *aratār*); *spārtt-* 'turn', mid. sg. 3 *sporttotār*, pl. 3 *sporttontār*, pple. *sporttomane* (= Turf. *sparcwa-trā*, *sparcwantrā*, *sparcwmām*). Like the *e*-verbs (Cl. III above), this class shows exclusively middle forms. Turfanian however shows a few active forms (e.g. sg. 3 *sparcwś-ām*).

It is even more difficult to make a judicious guess about the history of these verbs than about that of the *e*-verbs. One thinks immediately of IE *ā*. But *ā* in general seems to retain its quality and frequently also its quantity (see below, Classes V or VI). On the other hand it is possible that IE *ā* might become *o* under certain circumstances in Kuchean, as it does for example in *procer* = Turf. *pracar* = Lat. *frāter* 'brother'. This is not the time or place for detailed investigation of the problem.

Kuch. Cl. V, the *ā*-class, coincides in its numbering with SSS Cl. V. However,

of the dozen verbs listed here by Krause, only three have Turfanian cognates in the *ā*-class there. Of the rest, one has a cognate in Turf. Cl. III, one in Cl. VII, one in Cl. I, and the remainder are without any cognates at all. Examples are: (1) *lāk*- 'see', mid. sg. 3 *lkātār*, pl. 3 *lkāntār* (Turf. *lkātār*, *lkāntār*). The active forms *lakau*, *lkāt*, *lakaṃ* are found only as subjunctives (as opposed to Turfanian, where they are indicatives). The more usual present is in *-āsk*: *lkāskau* etc. (cf. below). (2) *šū*- 'eat', act. sg. 2 *šuwat* and *šwāt*, 3 *šūwam*, pl. 3 *šūwam*, *šwām-ne*, etc. (Turf. sg. 1 *šwām*, 2 *šwāt*, 3 *šwāš*, pl. 3 *šweñc*). This formation is, as Krause states, to be compared with the primary verbs of the type of Lat. *domāre*, *secāre*, etc.

Kuch. Cl. VI contains verbs in *-na*, *-ana* and *-nā*, *-anā*. The variation between these forms of the suffix is quite free; that is, I can find no reason to attach preference for one form or the other to ultimate, penultimate, or antepenultimate position. Thus, *kārs*- 'know' has sg. 3 *kārsanam* and *kārsnam*, *kārst*- 'cut off' has sg. 2 *kārsnāt* and sg. 3 *karsnam* or (with suffixed pronoun) *kārsnān-me*. In the middle we find sg. 3 *kārsanatrā* to *kārs*-, *kārsnātrā* to *kārst*-, *krapnatār* to *krap*- 'collect', *pālskanātār* to *pālsk*- 'think'.

In Turfanian we find, as the regular correspondents, two distinct classes, the *-na* (SSS VI) and the *nā* (SSS VIIa), e.g. Kuch. *skai*- 'strive', pl. 3 *skainam*, mid. pple. *skaināmane*, vbl. adj. *skainālle* : Turf. *ske*-, sg. 3 *skenas*, pl. 3 *skeneñc*, mid. pple. *skenmām*; Kuch. *wārp*- 'enjoy', mid. sg. 3 *wārpnatār* or *wārpanatār*, pl. 3 *wārpanantār*, mid. pple. *wārpanamane*, vbl. adj. *wārpanalle* : Turf. mid. sg. 3 *wārpnatār*, pl. 3 *wārpnanātār*, mid. pple. *wārpnamām*, etc.

Here belong also some of the correspondents to SSS VIIb (infix *ān*-class), e.g. Kuch. *kātk*- 'pass over', sg. 3 *kātkanam*, *kātksam* : Turf. pl. 3 *ktānkeñc*, inf. *ktānkātsi*. Other correspondents of the latter class, however, maintain their nasal infix (so alternate forms of *kātk*-) and are classed by Krause as VII. Not only does this class contain the IE *nā/nə* formation, as Krause says, but probably also remnants of the IE thematic *ne/no* presents—but without a trace of palatalization, though *ne* should appear as *ñā* in the sg. 2, 3, and pl. 2. A particularly clear instance would seem to be that of *tāl*- 'lift', sg. 3 *tallam* (*tāllam*, *tāllān-me*) with *ln > ll*, cf. Lat. *tollō*, *tollere* with the same development.

Kuch. Cl. VII is the nasal infix class. Verbs here are of two sorts: (1) roots ending in two consonants (eight in *tk*, one in *lk*, one in *rs*) with nasal and supporting vowel *ā* or *a* inserted between them, e.g. *kātk*- 'pass over' (also Cl. VI), sg. 3 *kāttānkām*; *kutk*- 'embody', mid. pple. *kutānkmane*; *pārs*- 'sprinkle', sg. 3 *prantsām*; *salk*- 'pull out', mid. sg. 3 *slānktār*; (2) the single root *pik*- 'write', sg. 3 *pinkām*, mid. pple. *pinkemane*. The inflection of both types is the 'thematic' but without palatalization, thus *pinkām*, not **piñcām* as one would expect (cf. Lat. *pingō*). The first type corresponds to SSS Cl. VIIb, with the difference that the Turfanian forms have a present stem in *ā*, e.g. *kāttānkām* : Turf. inf. *ktān-kātsi*.

Kuch. Cl. VIII is the *s*-class. In Kuchean the verbs in *s* (*š* when palatalized) are kept distinct as a class from formations in *sk* (*šš* when palatalized), while in Turfanian they are confused. That the same confusion was already beginning to take place in Kuchean is shown by the fact that a considerable number of the simple *s*-class have causative value. Examples are VIIIa (non-causative): *er*-

'cause, bring about', act. sg. 3 *eršām*, pl. 3 *eršem*, mid. sg. 3 *erštār*, pl. 3 *ersentār*, etc. (= Turf. *ar-*, act. sg. 3 *arāš*, pl. 3 *aršeñc*, mid. sg. 3 *arāštār*, etc.); VIIIb (causative): *pālk-* 'burn' (intr., primary pres. mid. sg. 3 *pālketār-ne*), caus. 'burn (trans.), torment', act. sg. 3 *pālkšām*, mid. sg. 3 *pālkštār*, cf. Turf. primary (?) pres. act. *pālk(s)e(ñc)*, pple. pres. *pā[l]kšant*, inf. *pālkāssi*; similarly Kuch. *plāk-* 'be in agreement' (primary pres. mid. pl. 3 *plākāntār*), caus. 'agree with, ask permission', impf. pl. 3 *plakšiyem*, cf. Turf. pret. mid. sg. 3 *plāksāt*.

At least one of the 'primary' *s*-verbs is in origin clearly a *sko*-formation, viz. Kuch. *prek-* 'ask', sg. 1 *preksau*, 3 *prekšām*, Turf. act. sg. 1 *praksa(m)*, mid. *prakšmār*, etc. (cf. Lat. *poscō*, Skt. *prcchati*). Since *kak* gave *ks* in this instance, it is possible that many of the other verbs in *ks*, primary or causative, may be of similar derivation.

The actual *sk/š(š)*-verbs are divided into three classes, IX, X, and XI, according to the entire form of the suffix; thus, Cl. IX covers the simple *sk*, *āsk*, *ask*, and *āsk*, Cl. X covers *nāsk*, *nask*, and *nāsk*, Cl. XI covers *sāsk* and *sask*. All three classes contain both primary and causative formations. A brief sampling survey follows.

Kuch. Cl. IX primary: (*sk*) *ai-* 'give', sg. 1 *aikau*, 3 *aiššām* (cf. Turf. *esam*, *ešt*, *eš*); (*āsk*) *nān-* 'show oneself', sg. 3 *nānāššām-ne*; (*ask*) *enk-* 'seize', mid. sg. 3 *enkastār*, pl. 3 *enkaskentār* (cf. Turf. *emtsāštār*, *emtsantrā*, etc.); (*āsk*) *kālp-* 'get', sg. 1 *kālpāsk(au)*, 3 *kālpāššām*, pl. 3 *kālpāškem* (but Turf. *kālpnātār*, etc., SSS Cl. VIIa); *lāk-* 'see', *lāskau*, *lākššām*, *lāškem*, etc. (but also pres. mid. *lkātār*, etc. = Turf. *lkām* etc., act. and mid. both, SSS Cl. V; cf. above).

Most of the causative verbs listed here show the suffix form *-āsk*, rarely *-ask*. Initial palatalization occurs in a minority of verbs. Krause subdivides these causative verbs according to the formation of the subjunctive and/or the preterit, as follows:

(a) The *sk/šš* suffix is found also in the subjunctive, but not in the preterit. The indicative present and subjunctive are thus formally identical (and so are also then the imperfect and optative): *kārs-* 'know' (prim. pres. sg. 3 *kārsanam*), caus. pres. sg. 3 with initial palatalization *šarsāššām*, opt. *šarsāšši*, but caus. pret. *šārsa*; *prutk-* 'be enclosed' (prim. pres. mid. sg. 3 *prutketār*), caus. pres *prutkaskau*, *prutkāššām* (*prutkaššām*), caus. pret. *prautka*.

(b) The *sk/šš* suffix not only is found in the subjunctive, but has been extended to the preterit also: *kārp-* 'descend' (prim. pres. *korpotār*), caus. pres. mid. sg. 3 *kārpastrā*, opt. act. sg. 1 *kārpāššim*, pret. *karpāššasta*; *šau-* 'live' (prim. pres. pl. 3 *šawem*), caus. pres. *šawaššām*, pret. *šawšate*; *lāk-* 'see' (prim. pres. *lāskau*), caus. pres. pple. *lakāskemane*, pret. *lakāšša*.

(c) The *sk/šš* suffix is found only in the present. The subjunctive and preterit are formed from a 'strong' grade of the root. Only one example is actually attested here: *kān-* 'come into being' (prim. subj. *knetār*), caus. pres. *knastār*, subj. *kyānamar*, pret. *kyāna*. While it is probable that the verb *lūk-* belongs here too, the divergences between Kuchean and Turfanian are too great for us to posit it as a fact on the basis of the latter dialect; the verb had better be listed with the numerous causatives of this class (93-5) for which there is not enough evidence to classify them in any group.

Kuch. Cl. X consists of the verbs in *nāsk*, *nask*, and *nāsk*. In contrast to Class

IX, the primary verbs here are either in *nask* or *nāsk*, while the causatives (as also regularly in IX) show the medial vowel *ā*. Primary: (*nask*) *au-* 'meet', act. sg. 3 *aunaṣṣan-me*; *ri-* 'leave', mid. sg. 3 *rinastār*, etc. It is interesting to note that all verbal roots ending in vowels, like the two examples just given, have extended the *n* throughout the inflection, e.g. (to *au-*) subj. mid. pl. 3 *aunantār*, pret. *auntsante*, pple. pret. *aunu*, etc.; not so however the roots in a consonant, e.g. *kām-* 'come', act. sg. 3 *kānmaṣṣām* (regular metathesis of *mn* > *nm*), also in subj. *śanmām*, but pret. *śem*, pple. pret. *kekamu*, etc. Primary verbs in *nāsk* are typified by *kār-* (*kāry-*) 'trade', mid. sg. 3 *kārñastār* (pret. *kāryātai*, pple. *kāryoṣ*), or by *pāk-* 'intend', mid. sg. 2 *pāknāstar* (pret. *pkāte*). The etymology of the former clearly shows that the suffix is an extension in *sk* of the IE *nā/nə* formation (cf. Skt. *krīṇāti* 'buys', OIr. *crenaim* 'I buy').

As for the causatives in *nāsk* (*nask*), they are made from primary verbs in *nask* or *nāsk* (e.g. *tanmāṣṣām* 'begets' to primary *tānmastār* 'is born'), except for *śanmāstār* 'binds', which Krause considers a causative, though no primary form is attested. The forms of *stāl-* are too uncertain to be useful here. In Turfanian we find the corresponding suffixes *nās* (SSS VIII) and *nās* (X), e.g. sg. 3 *kumnāṣ* (: Kuch. *kānmaṣṣām* above), *pāknāstār* (: Kuch. *pāknāstar* above). There seem to be no causatives of similar formation in Turfanian.

Kuch. Cl. XI consists of five verbs in *sask* (*saṣṣ*) and *sāsk* (*sāṣṣ*). Of these, three are primary: *āks-* 'announce' (1 sg. *aksaskau*), *āks-* 'awaken' (pple. act. gen. pl. *āksaṣṣeñcantse*), and *auk-* 'grow' (3 sg. act. *auksāṣṣām*). Two are causative: *tu-*, *twā-* 'kindle' (3 sg. mid. *twasastār*, prim. pres. 3 sg. act. *twāṣṣām*) and *su-*, *swā-* 'rain' (1 sg. act. *swāsāskau*, prim. pres. 3 sg. act. *suwaṃ*). The nearest Turfanian parallel would be the two verbs in *sis*: *āks-* 'teach, announce' (sg. 1 *āksisam*, 3 *āksiṣ*) and *oks-* 'grow' (sg. 3 *oksiṣ*, mid. pple *oksismām*). In fact the identity of the two suffixes is affirmed by the identity of the two verbs for 'teach, announce' and 'grow'.

Finally there is Kuch. Cl. XII, the verbs in *ññ* (with connecting vowels: *aññ*, *aññ*, *eññ*). These are to be identified, of course, with the Turfanian Class XII in *iññ*. The formation includes both clearly denominative verbs, as also in Turfanian (Kuch. impf. *tāñkwaññet* : *tāñkw-* 'love', cf. Turf. inf. *tuñkiñtsi* : *tuñk-* 'love'), and what seem to be primary derivatives (Kuch. *kāsk-* 'scatter', impf. mid. *kāskaññitār*; *klānts-* 'sleep', impf. *klāntsaññi*, but Turf. *klis-*, pres. *klisnāṣ* VIIb). In so far as these are denominatives they must have their origin in derivatives of *n*-stem nouns (type Gk. *ὄνομαλιν*). It is well known what an important role the *n*-stems play in the Tocharian declension.⁶

I shall not dwell in detail here on the Kuchean imperfect (Krause 103-9), since I have discussed it in some detail in the above-mentioned article on the Tocharian imperfect and preterit. In any case, on the general origins of the Kuchean imperfect there seems to be common agreement. It reflects in some fashion the PIE optative, either thematic or athematic or both. As Krause notes, the Kuchean imperfect is always derived from the present stem, a rule which contrasts sharply with the Turfanian situation, where, as I believe I have shown

⁶ Cf. Pedersen, *Tocharisch vom Gesichtspunkt der indoeuropäischen Sprachvergleichung* 40, 63-ff.; Couvreur, *Hoofdzaken van de tochaarse klank- en vormleer* 43 ff.

in the same article, the imperfect is largely of the same origin as the preterit. The sole exceptions to these statements are the verbs 'be' and 'go'.

The subjunctive, however, will hold our attention a bit longer. In Kuchean it is to be related to the present indicative system, so far as it is true that the greater share of the stem formatives are common to both moods. In fact, only two subjunctive formations (those in *i* and *ñ*) are not found also in the present system, and only three present formations (*o*, *n*-infix, and *s*) do not occur also in the subjunctive. As a result, while for the most part any particular present indicative system will contrast with the subjunctive of the same verb, in many instances the forms are identical. So far as subjunctives to pres. Cl. I are attested, in fact, they are mostly identical with the present. Only one verb shows a contrast: *sanāp*- 'anoint', pres. *sonoptār* : inf. *sanāpatsi*. On the other hand, the subj. Cl. I is used in a wide variety of present stems (cf. the tabulation in Krause 118, 119.7). In pres. Cl. II, a considerably smaller percentage of the roots show the same formation in the subjunctive (only 11 out of some 28 verbs, as opposed to 10 out of some 15). In contrast to these first two present classes, presents in *e* (Cl. III) never show *e* subjunctives, and whereas *e*-presents are middle, *e*-subjunctives have active endings. There are of course no *i*-presents to compare with *i*-subjunctives (Cl. IV).

Subjunctive Cl. V, the *ā*-subjunctive, is the dominant formation, occurring in well over a hundred verbs, of which ten form *ā*-presents also; e.g. sg. 1 *prāskau*, 2 *prāskat*, 3 *prāskam*, pl. 3 *parskam* 'fear', where at least *prāskau* is both present indicative and subjunctive (cf. Krause 259); sg. 2 *śuwat* or *śwāt*, 3 *śūwam*, pl. 3 *śūwam* or *śwām* 'eat' (Krause 294). However, presents and subjunctives of the *ā* class are not always identical in form, e.g. pres. mid. sg. 3 *peññatrā* 'stretches', subj. act. sg. *pānnam*. Here we see clearly that the present system was once quite different, in all probability an original *yo*-verb (cf. also Turf. sg. 3 *pañwāṣ*).

Four verbs form their subjunctive stem in *nā* (Cl. VI). Three of these have their presents in *nāsk* (pres. Cl. X), one in *āsk* (IX). It is noteworthy, I believe, that the latter verb, *kālp*-, has a *nā*-present in Turfanian. That is, the Kuchean subjunctive and the Turfanian present are identical in formation and hence probably also in origin.

The *ñ*-subjunctive (Cl. VII, 140 f.) is cited for only two verbs: *lāt*- 'go out' (pres. sg. 1 *lnaskau*, 3 *lnaṣṣām*) and *we*- 'say' (pres. sg. 1 *weskau*, 3 *weṣṣām*). For the former verb, however, the evidence is only indirect; decisive subjunctive forms are absent, unless 1 pl. *lāññam* is actually subjunctive and not optative (for **lāññem* as Krause himself prefers). Clearly subjunctive are only sg. 2 *lat* and 3 *lam*; and of these, *lat* must be (an error?) for **lant* (for **la-n-t-ñ-t*, just as *lam* must be from **la-n-t-ñ-n*). Optative forms such as sg. 2 *lyñit*, 3 *lāñi* or *lāññi*, of course, furnish no evidence at all, since the *n* would become *ñ* in any case. For *we*-, however, the *ñ*-subjunctive is made clear by sg. 1 *weñau*, pl. 3 *weñem*.

From the comparative point of view it is noteworthy that on the one hand, as Krause points out, the *ñ*-subjunctive in Turfanian belongs likewise to many *sk*-presents (part of them labelled causative in SSS), but that on the other there is no connection at all in Kuchean between the *s*-preterit and the *ñ*-subjunctive as

there is in Turfanian. The preterit of *lāt-* is a 'thematic aorist' in Kuchean: sg. 1 *latau*, 2 *lāt*, 3 *lāc* or *lac* < **letet*, pl. 3 *lateṃ* < **letont*,⁷ and possibly also originally in Turfanian, e.g. sg. 3 *lāc* but extended analogically to *lā-* in bound forms. The third plural Turf. *lār*, however, would seem to belong to an *s*-preterit. The Turfanian subjunctive sg. 3 *lāñcāṣ*, pl. 3 *lāñceñc* appears to be in origin a 'thematic' present indicative (like *yāmāṣ*, SSS 376). This type is, as we have seen, widespread in Kuchean (subj. = indic. Cl. II, Krause 122 ff.). To this same type may also be attributed, from a purely descriptive point of view, Turf. *weñāṣ* (subj. to *trānkām* = Kuch. *weṣkau* 'speak'), but here the subjunctive stem is identical with the preterit; e.g. sg. 1 *weñā*, 3 *we*, in bound forms *weñā*, etc. = Kuch. pret. sg. 1 *weñāwa*, 3 *weña*, etc.

As subjunctive Cl. IX are considered those forms in *-sk-* (*-ṣṣ-*) which, though identical with the present indicative, are adjudged on syntactical grounds or on the basis of a Sanskrit original to be subjunctive. There are here noncausative formations in both *-āsk-* (*an-āsk-* 'breathe', *al-āsk-* 'be sick', etc., five in all) and *-ask-* (rarely *-ask-*). As is obvious, the choice between indicative and subjunctive use of identical forms leaves the way open for great uncertainty in classification. The decision on the basis of the Sanskrit form is more trustworthy but not absolute. For descriptive grammar the question here, as elsewhere where present indicative and subjunctive fall together, is of no significance.

Unlike the presents in *-āsk-* and *-ask-* (*-ask-*) above, present indicatives of Cl. X (*-nāsk-*, *-nāsk-*, *-nask-*) and Cl. XI (*-sāsk-*, *-sask-*) form a subjunctive identical with the indicative only in causatives. Primary verbs of both formations have subjunctives without *-sk-* (*-ṣṣ-*). Examples: Cl. X *tām-* caus. act. 'beget', indic. or subj. sg. 3 *tanmāṣṣām*, pl. 3 *tanmāskem*, inf. (from subj. stem) *tanmāstsi*; but prim. mid. 'be born', indic. sg. 3 *tānmastār*, subj. (Cl. III, *e*-subj.) sg. 3 *cmetār*; Cl. XI *āks-* 'announce', prim. indic. sg. 3 *āksaṣṣām*, subj. sg. 3 *ākṣām* (thematic!).

The *ññ*-subjunctive (Cl. XII) is attested, according to Krause (146-7), for only five verbs, three primary and two denominative. Of the former, two (*ārc-* 'must' and *tser-* 'deceive') are probably also *ññ*-presents. One, *suk-* 'hang down' (opt. *sukaññiyentrā*), is clearly not (pres. sg. 3 *suknaṃ* (Cl. VI). The two denominatives in *ññ* for which *ññ*-subjunctives are assumed (on scanty evidence) are *tāñkw-aññ-* 'love' and *skw-aññ-* 'feel happy'. On the other hand, presents in *ññ* in general, both primary and denominative, do not necessarily show *ññ*-subjunctives (cf. Krause §142).

A considerable amount of space has been devoted here to a discussion of the subjunctive as presented by Krause, in the hope of bringing out not only its heterogeneity but also the lack of co-ordination between indicative and subjunctive formations. This great confusion of forms is in sharp contrast to Turfanian, where the majority of subjunctives are derived from the preterit stem. To be sure, we find Turfanian subjunctives characterized by particular stem-building suffixes independent of the preterit, of which the counterparts are to be sought in Kuchean; but they give the impression of being the irregular remnants

⁷ The final *t* of the root can of course represent any IE dental. On the etymology and the identification as a thematic aorist, see Pedersen, *Tocharisch* 189 bottom.

of an older state of affairs which was giving way to a new system as rapidly as possible. Certainly the history of the subjunctive will be one of the hardest chapters to write in any future comparative Tocharian grammar. Possibly its great diversity and the actual widespread identity of indicative and subjunctive forms in Kuchean is another bit of evidence in favor of the hypothesis, current in several quarters, that the subjunctive is of very late origin in the Indo-European family—an assumption that would explain its great diversity in most languages and its complete absence in some (Germanic, Balto-Slavic, Hittite).⁸

The section on the imperative (147-53) is exceedingly well done and very detailed. It was precisely the imperative of Turfanian which received the least attention in SSS. In a way, the account here for Kuchean seems to compensate for the neglect in the earlier instance. The formation of the imperative again forcibly reminds us of the divergence of the two dialects: whereas the Turfanian imperative is based regularly on the preterit stem (cf. SSS 325), the Kuchean imperative is derived from either the subjunctive or the preterit stem, but as regards the suffixes used there is close agreement with those of the preterit. Here again we seem to see the extension and regularization in Turfanian of a state of affairs still only incipient in Kuchean. The same comment applies also to the fact that the prefix *p-* (*pā-*, *pu-*), which is obligatory in Turfanian, is frequently omitted in Kuchean, especially before *p-* but also elsewhere.

A form peculiar to Kuchean is the so-called durative (153). Only five forms (from four verbs) are attested. Its formation is simple: the imperfect sign *ʔ* is added to the preterit stem, which then takes preterit endings. The imperfect, by contrast, is based on the present stem and has present endings (except for act. sg. 1 and 3). Compare pres. act. sg. 3 *šuwam*, impf. *šuwoy*, with pret. act. sg. 3 *šūwa*, pl. 3 *šawāre*, dur. sg. 3 *šawiya*, all from *šu-*, *šwa-* 'eat'. This formation I should count as an innovation in Kuchean. I can hardly believe that it goes back to an Indo-European aorist (or perfect) optative.

The preterit, like the imperfect, needs only a brief comment here concerning Krause's disposition of the materials; the above-mentioned article in the Edgerton number treats it in great detail. In general Krause's classification follows as far as possible that used by SSS for Turfanian. However, Turfanian Ib ('ablauting' pret.) appears as Kuchean Ia (e.g. Turf. act. sg. 3 *kčāk*, pl. 3 *katkar* : Kuch. sg. 3 *šatka*, pple. *kātkau*; Turf. mid. sg. 3 *kālpāt*, pl. 3 *kālpānt* : Kuch. act. sg. 3 *kalpa*, pl. 3 *kālpāre*); and Turfanian Ia ('strong vowel' pret.) as Kuchean Ib (e.g. Turf. act. sg. 3 *tāk*, pl. *tākar* : Kuch. *tāka*, *takāre*; Turf. mid. sg. 3 *kropat*, pl. *kropant* : Kuch. *kraupāte*, *kraupānte*). There are of course very many shifts, in particular some Turfanian Cl. III (s-preterits) appear here: Turf. *māškās* : Kuch. *maska*; Turf. *asās* : Kuch. *asāre*. And here also are placed two verbs in Turfanian Cl. IVa: act. sg. 3 *klā* : Kuch. *klāya*, and *wā* : Kuch. *wāya*, from Turf. *klā-*, Kuch. *klāy-* 'fall', as well as the suppletive root *wā-* to pres. Turf. = Kuch. *āk-* 'lead'. On the other hand Turf. *lyā-* 'wipe, sweep' (abs. *lyālyorāš*, vbs. *lyā-lune*), classed as IVa in SSS, shows an s-preterit in Kuchean; sg. 3 *lyyāsa*.

The Turfanian reduplicated preterit, SSS Cl. II, is, as usual, equated to the Kuchean 'long vowel' or 'strong vowel' type of *cāla* : Turf. *cacāl* (to *tāl-* 'lift'),

⁸ Cf. Pedersen 191 f. I generally subscribe to his views on this matter.

šārša : Turf. *šašārs* (caus. to *kārs*- 'know'), *raittate* (mid.) : Turf. *rarity* (act., caus. to Kuch. *ritt*-, Turf. *ritw*- 'be bound'). The class in both dialects is largely causative in value. The equation of the two, and the derivation of the Kuchean unreduplicated forms from originally reduplicated forms which are maintained in Turfanian, go back to W. Schulze.⁹ My own dissent from this view is stated in my article on the Tocharian imperfect and preterit, and need not be argued here.

Kuchean and Turfanian Cl. III preterits, the so-called s-preterits, coincide exactly; cf. Kuch. sg. 3 *neksa* : Turf. *ñakās* (to *nāk*- prim. 'disappear', caus. 'destroy'). The further identity of the s-preterit with Turfanian imperfect Cl. III is defended in the above-mentioned article. One striking difference between the two dialects with regard to the s-aorist, however, is that Turfanian shows two different middle formations, one with s throughout, the other (attested only in the 3rd sg. and pl.) without s, e.g. act. sg. 3 *ñakās* but mid. sg. 3 *nakāt*, pl. 3 *nakānt*, as opposed to act. *prakās*, mid. *prāksāt*, *prāksānt* (to *pār*k- 'ask'), while Kuchean has both act. *neksa*, mid. *neksate*, and act. *preksa*, mid. *park-sante*. This is one of the few instances where the verbal system of Turfanian gives at least superficially the impression of greater complexity and less regulation by forces of analogy than the more archaic Kuchean.

The Kuchean §§ preterit (Krause Cl. IV) is formed from three primary verbs with presents in *sk* (§§), and from a score or so of causatives of present classes IX–XIb. Examples, primary: pres. sg. 1 *wināskau* 'honor', pret. sg. 3 *wināšša-me*; pres. sg. 1 *yamaskau* 'do', pret. sg. 1 *yamāššāwa*; caus.: *kārp*- 'descend', pres. mid. sg. 3 *kārpastrā*, pret. sg. 3 *karpāššate*; *soy*- 'be satisfied', pres. act. sg. 3 *soyāššām*, pret. sg. 3 *soyša*. Krause remarks (180, §183 Anm.) that this formation is found in Turfanian only in pret. participles (and absolutes) in *-(š)u* such as *wawim-surāš* (he fails to cite *āršu* here and in §185). Of course this is true so far as the preterit is concerned. It appears in Turfanian rather as the 'regular' imperfect of all causative verbs.¹⁰

Krause's Cl. V (SSS IVb), 'Präteritum mit sonstigen Suffixen', includes *we*- 'speak' (*weñāwa*, *weñāsta*, *weña* = Turf. *weñā*, *weñāšt*, *weñā*-) and forms in *ññ*. But whereas Kuchean has only three examples of the latter type, probably all from denominative presents, Turfanian shows a considerable extension of the formation, in particular to the two verbs *āks*- 'announce' and *oks*- 'grow' with present stems in *-is*- (*āksisam* etc.). This is another example of the Turfanian tendency to distinguish its verbal stems with greater contrast than we find in Kuchean. In a way one is reminded of the confusion of the Irish verbal system over against the clarity of the Welsh.

Krause's Cl. VI preterits are those of the verbal roots *kām*- 'come' and *lānt*- (or *lāt*-) 'go forth'. The same two preterits are classed by SSS as V, since the Kuchean §§ preterit had to be intercalated as IV. In Kuchean, as opposed to Turfanian, the preterit of *kām*- (PIE **g^{em}-*) is well attested: act. sg. 3 *šem*, pl. 1 *kmem*, 3 *kameṃ* (*kmeṃ-ne*); mid. sg. 2 *kamtsatai*, pl. 3 *kamtsante*, pple. *kekamu*, *kekmu*. In Turfanian we find only the pple. *kakmu*. The appearance of *e* in *šem*

⁹ SBAW 1924.166 ff. = *Kleine Schriften* 239 ff.

¹⁰ See my article in the Edgerton number of *LANGUAGE*.

is strange. If this is an athematic aorist, as Pedersen concludes,¹¹ to be equated to Skt. *agan* (**e-g^wem-t*), I should expect *a* or *ā* in Kuchean rather than an *e*. Moreover, we actually do find IE *e* represented by *a* and *ā* in the preterit of *lā(n)t-* (act. sg. 1 *latau*, 2 *lāt*, 3 *lac*, pl. 2 *latso*, 3 *latem* : Turf. act. sg. 1 *lcā*, 3 *lāc*, *lcā-*, pl. 3 *lcār*), where we are dealing in all probability with an original thematic aorist (note the palatalization in the third sg. but not in the third pl.). I see no immediate explanation.¹²

The fifth of the unequal divisions of the book gives a survey of the Kuchean verbal endings, with the corresponding Turfanian forms placed after them in brackets. The remarkable contrasts between the two dialects in this respect are thus placed in bold relief; e.g. pres. (and subj.) act. sg. 1 *-u* : *-m*, 3 *-ṃ* : *-ṣ*, pl. 1 *-m* : *-mās*, 2 *-cer* : *-c*, 3 *-ṃ* : *-ñc*; opt. sg. 3 *-* : *-ṣ*; pret. mid. sg. 1 *-mai* : *-e*, *-we*, and others less striking.

In this section is found the only serious excursus into comparative grammar: Zur historischen Erklärung der Personalendungen (199–203). A few comments are perhaps in order. I am happy to note that the older view that the Kuchean ending *-u* (*-au*, = *eu*) could be equated to the Turfanian *-m* is here abandoned.¹³ Turf. *-m*, as has been generally accepted, is of course PIE athematic *-mi*. Whether the Kuchean *-u* can go back to the thematic *-ō* is still problematical, but I am delighted that Krause does not completely reject the possibility that the latter may be involved here.¹⁴ Krause's view that the sg. 2 *-t* (of both dialects) comes from PIE *-tha* (Skt. *-tha*, Gk. *-θα*) is certainly a simple explanation, phonetically, but the question remains: why a perfect ending in the present? Similarly the explanation of the sg. 3 Kuch. *-ṃ* (= *n*, of course) and Turf. *-ṣ* as the secondary pl. 3 IE *-nt* and the primary sg. 2 IE *-si* respectively, though phonetically acceptable, seems to me to involve a lot of inexplicable juggling of forms (in spite of the English sg. 3 in *-s* and the Scandinavian in *r* < *s*). The distinction between the Kuchean pl. 1 *-m* and the Turfanian *-mās* is explained by calling upon both IE primary *-mes* and *-mesi* (Ved. *-masi*). Kuch. *-m* can of course just as well represent the secondary *-me*; and the absence of palatalization of final *-s* in Turf. *-mās* (we would expect *-māṣ*) makes the Vedic parallel doubtful (unless from IE *-mesā*). There seems little doubt that the Kuch. pl. 2 *-cer* is to be analyzed as *-c + er*, where the *-c* = Turf. *-c* from IE *-te*, but what the *-er* is must remain uncertain. Like Krause, I think the transfer of the middle ending *-r* is extremely improbable. The pl. 3 Kuch. *-ṃ* and Turf. *-ñc* are clearly to be derived—as Krause (and before him Pedersen) derives them—from the PIE secondary *-nt* and the primary *-nti*.¹⁵

¹¹ *Tocharisch* 185.

¹² Pedersen, *Tocharisch* 219, considers that PIE *e* appears normally as Toch. *e* (thus falling together with PIE *o*), and that the change to *a* is not the rule: 'Eine Tendenz, *e* in *a* zu verwandeln, findet sich in beiden Dialekten, in A jedoch in viel grösserer Ausdehnung als in B' (220, §104).

¹³ For the theory that *-m* > *-n*, see Couvreur, *Hoofdzaken* 49, 55. The theory is rejected by Pedersen, *Tocharisch* 141.

¹⁴ Cf. Pedersen's conjectures, loc.cit.

¹⁵ Cf. op.cit. 144. Couvreur, on the other hand, insists that Kuch. *-ṃ* must represent IE *-nti* also, op.cit. 56.

As for the endings of the perfect active and of the entire middle, a competent discussion of Krause's derivations and suggestions (many of them excellent) would unduly prolong this already overextended review. Similarly for the non-finite forms. Just one note here. The act. pres. pple. of originally thematic verbs is in *-ēñca*, with palatalization of the preceding consonant, as in *aśēñca* (: *āk-* 'lead') = Lat. *agent-*. This shows, of course, that the original vowel was *e* (not *o* as in Greek and Gothic). However, I should expect here **aśañca* (cf. sg. 3 *aśām*, *aśan-me* < **age[t]* as opposed to pl. 3 *ākem*, *aken-me* < **agont*), unless indeed the retention of *e*-quality is due to the following palatal consonant cluster. The middle pple. *ākemane* < **agomen-os* is of course 'regelrecht'.

A final chapter of this section is devoted to the suffixed personal pronouns. The excuse (if one is needed) for including a discussion of pronouns in a work devoted to the verb, is that these pronominal forms had become so closely assimilated to the verbs to which they were attached that the whole was treated (and no doubt was felt) as a unit. In fact the alternation of 'thematic' *ā/a* (< IE *e*) in verb forms before the suffixed sg. 3 *-ne* and pl. 1, 2, 3 *-me* is one of the strongest pieces of evidence that the accent fell on the penult, e.g. *aśām* 'leads' : *aśan-me*; *campām* 'can' : *cāmpan-me*; *palkām* 'shines' : *palkān-ne*. The closest parallel in any Indo-European language would be in Old Irish. From the comparative point of view, longer forms of the sg. 1 and 2 in Turfanian are peculiar: *-ñi* and *-ci*, as opposed to Kuchean *-ñ* and *-c*. The sg. 3 *-ñ* (= *n*) and pl. *-m* in Turfanian, beside the longer Kuchean forms *-ne* and *-me*, are what we expect. Perhaps the Turfanian *-ñi* and *-ci* have been extended on the analogy of the genitives *ñi* (masc.), *nāñi* (fem.), and *tñi*, since these pronouns are very commonly used in a genitive sense (cf. SSS 274 f., Krause 207 ff.). That the Turfanian *-ñi* and *ci* are not at all original is further indicated by the fact that they alone (as opposed to *-ñ* and *-m*) are appended to nonfinite forms (participles and infinitives) and even to substantives in nominal sentences (SSS loc.cit.). In Kuchean the more original *-ñ* and *-c* (like *-ne* and *-me*) may be appended only to finite verb forms (Krause 206).

The following section is one which has always been lacking for the beginner in Turfanian: a set of complete verbal paradigms, with assumed forms where none are attested. The author has marked as assumed only those forms which are not attested for a particular tense formation; forms that can be assumed, in the opinion of the author, only provisionally or not at all, are identified by a following query or are omitted entirely. In general, conservatism is the rule.

For me the most valuable part of the book is the Verbalverzeichnis, containing the forms of both the edited and the unedited Berlin collection; and of the Hoernle collection (likewise both edited and unedited texts), as well as forms from certain unedited Paris texts. Here again we find an advantage for the beginner, or at least for the nonspecialist, which is lacking in the Verbalverzeichnis of Turfanian in SSS—namely, the careful designation of the class to which each tense belongs. Further, primary (noncausative) formations are marked by a boldface G (Grundverb) to distinguish them from the causatives (K). Otherwise the arrangement and terminology follow, with minor exceptions, the precedent set in SSS. It is to be noted that the author uses the traditional 'A'

to designate Turfanian equivalents, instead of 'Osttocharisch', as he does earlier in the book, in verbal listings. Is this possibly an indication that the brief, convenient A-B distinction, no matter how absurd, will eventually win out against Krause's efforts to substitute the more significant Ost- and West- and my own substitution of Turfanian and Kuchean? A small matter after all.

In conclusion I can only repeat a word from the first sentence of this review: indispensable. All references to the verbal forms of Kuchean (Westtocharisch, Tocharian B, as you will) will henceforth be based on Krause's volume. It will be the foundation of all future discussions just as all discussions of Turfanian rest on SSS. No doubt there will be changes made in the classificatory system. I myself would prefer, where possible, to retain the parallelism with the system already established for Turfanian; or perhaps—and that is worth trying—one should now work out a system better suited to both. Such an attempt will be the beginning of comparative Tocharian grammar, for which, at long last, we are now almost ready.

Grammaire de la langue védique. By LOUIS RENOU. (Collection 'Les langues du monde', Série grammaire, philologie, littérature: Volume 9.) Pp. 454. Lyon: IAC, 1952.

Reviewed by FRANKLIN EDGERTON, *Yale University*

This is a descriptive grammar of the Vedic mantras (sacred verses and formulas). It is historical only within its own frame of reference, noting linguistic changes first within the Rigveda (early and late parts), then between it and the Atharvaveda (these two are, of course, much the most prominent in the treatment, because the oldest—especially the Rigveda), and the still later Vedic works containing mantra material.

There can be no doubt of the propriety and desirability of the book's aim. The monumental *Altindische Grammatik* of the late Jacob Wackernagel, now being continued by Albert Debrunner, is only about five-eighths published; furthermore, like Whitney's *Sanskrit grammar*, it includes all stages of Vedic and Sanskrit together, and the language of the mantras unquestionably deserves separate treatment. It has already received such treatment in other works, some of great value, but none that seriously rival Renou's book, which has qualities that place it on a higher level than any earlier complete Vedic grammar.

A large part of Renou's very extensive previous publication seems preparatory to this task, and makes predictable his success in it. His work on Vedic grammar began with his important doctoral thesis (1925), *La valeur du parfait dans les hymnes védiques* (covering a wider range than the words suggest). In 1930 appeared his two-volume *Grammaire sanscrite*, which perhaps deserves to be called the only thoroughly independent and highly valuable grammar of classical Sanskrit that has appeared in the west, except Whitney's and Wackernagel's. There followed in rapid succession numerous articles and monographs containing penetrating studies on details of Vedic grammar and interpretation, as well as a number of stout books which helped to prepare for the present work: *Les maîtres de la philologie védique*, 1928; *Bibliographie védique*, 1931; *Monographies*

sanskrites, 1937; *Les écoles védiques et la formation du Véda*, 1947; and others. Also pertinent to this work, and almost as directly, were his searching studies in classical Sanskrit grammar, far too numerous to list here; e.g. the edition and translation of the grammatical work *Durghaṭavṛtti* (1942-), *Terminologie grammaticale du sanskrit* (3 vols., 1942), and a translation of Pāṇini with extracts from commentaries (1948-). In passing, it should be emphasized that Renou has published in many other fields of Indic studies, literary and cultural, some abstruse and difficult: among them Upaniṣads, kāvya, philosophy, religion, and geography. Many an Indologist might be content to rest his reputation on Renou's non-grammatical works alone. He is living proof of the fact that a really supreme grammarian should be also (not only a philologist, that is a text-interpreter, but) a profound knower of the whole culture carried by the language. In fact, Renou is a worthy representative of the glorious traditions of French Indology, in the line of Burnouf, Bergaigne, Senart, and Sylvain Lévi (to mention only the dead, and only the greatest among them).

Everyone who deals with the Veda will have to keep this book at his elbow, and should read it periodically from beginning to end, to his great enjoyment as well as professional profit. In the briefest possible form, but with little sacrifice of exhaustiveness to simplification, it summarizes all that can be said with reasonable confidence of the Vedic language, as it appears to one of its two or three greatest living masters. Even an intelligent beginner should, I think, find it profitable, and for the most part not difficult, to read and study it. There are no paradigmatic tables. The Preface says: 'Nous avons voulu insister plutôt sur les singularités. La nature de la documentation, les tendances du style védique sont ainsi faites que la tradition des *mantra* apparaît, pour une large part, comme un répertoire d' "anomalies", de tentatives grammaticales et stylistiques. Ce serait ne pas y être fidèle que de donner de cet ensemble une image facile et linéaire.' To this sound statement I would only add that Renou has been careful (and this is one of the book's most striking virtues) to inform his readers, unobtrusively but constantly, as to the scope and frequency or rarity of each item, down to isolated occurrences. He also helps the beginner, and even the advanced scholar, by regularly translating his examples. Many phrases and whole pādas, or even longer passages, are quoted and fully translated. The original sources are indicated by abbreviated titles (unidentified passages are from the Rigveda; I and X indicate its first and tenth books), but exact references are usually not given; they can almost always be easily found by any informed reader, and valuable space is thus saved.

More embarrassing, and rather regrettable, is the omission (no doubt to save space) of any bibliographical references whatever, in the body of the grammar. Experienced Vedists will, to be sure, not need to be told in many, perhaps most, cases which of the author's statements are controversial; and he himself frequently adds 'douteux' or the like. But beginners need more information; and this reviewer, at least, has sometimes found it hard to satisfy his curiosity as to the origin of a particular doctrine or theory, unfamiliar to him (a mediocre bibliographer, it must be confessed). Of course many of these, or certainly some, are no doubt both original and new.

The grammar proper occupies less than 400 rather small pages (32–35 lines per page; no footnotes, but many notes in smaller type inserted in, or after, the 466 numbered sections). Numerous cross-references, and an excellent index of Vedic words, help to save space without loss of clarity. The style skillfully combines brevity with lucidity. There is a four-page bibliography, a selection largely of recent publications not listed in Wackernagel–Debrunner.

In his Preface, Renou says that he has ‘suivi de près’ Wackernagel(–Debrunner) on the subjects of noun composition and noun inflection. He has however made important additions to that admirable work. Two examples: in §299 he records *paṣṭha*-(*vāh*, *paṣṭhauhī*) ‘fifth’ (ordinal numeral), convincingly established by him in *BSL* 43.38 ff. (1946). This very interesting form has always been misinterpreted, usually as Middle Indic for *prṣṭhá* ‘back’—a very old ‘popular etymology’, going back perhaps to late Vedic times (*MS* 3.11.11d, 158.9; but p.p. *paṣṭa*-, sic), and persisting at least to 1932 (*Vedic variants* 2.296, I regret to say). It seems to me clearly a riming adaptation of original *pakthá* (recorded at least as n.pr., ‘Quintus’) to match *ṣaṣṭhá* ‘sixth’. Renou, in *BSL*, thinks of this (with other less likely theories), but in the present work offers no explanation. — In §282 he recognizes that *tmán* and *tán* ‘self’ seem to form a paradigm’ (better, to be automatic variants of one another; cf. *Lg.* 19.116). Much earlier, Ludwig (*Rigveda* 5.352–3) saw the ‘essential identity’ of the two, but no one heeded him; the time was not then ripe for perception of the phonemic explanation. It might, I think, have been added that originally (and still in RV with almost no exceptions) *ātmán* ‘soul, self’ formed ‘strong’ cases only, *tmán*/*tán* ‘weak’ cases only, all constituting a single paradigm (*Lg.* loc.cit.).

The grammar is divided into six chapters, as follows: 1. Phonétique (11–111): alphabet, pronunciation, vocalism, consonantism, accent, word-final, *saṁdhi* of vowels and consonants; 2. Formation du nom (112–86): composition, primary and secondary derivation (suffixes); 3. Flexion du nom (incl. pronoun and numeral; 187–245); 4. Le verbe (246–314); 5. Les invariants (315–32): preverbs, prepositions, adverbs with adverbial compounds and derivatives, interjections (conjunctions are listed and discussed in the appropriate parts of the next chapter); 6. Syntaxe (333–403): order of words, agreement, pronouns, noun cases, nominal clauses, verbal nouns, voice, tense and mood, causative, coordination, negation, interrogation, subordination, relative, ‘complétive’ (chiefly clauses beginning *yád* ...), ‘circonstanciellles’ (*yád* ... etc.), final-consecutive, conditional, direct discourse, style (ten pages, including some very perceptive and useful observations).

In a field so obscure as the Veda, so full of difficult problems, some of which no doubt will never be solved and for many of which equally distinguished scholars have arrived at radically different solutions, it goes without saying that any serious student will have some more or less different views from any other, or at least would prefer to use different language on some points. The relatively few matters on which I might venture to differ from Renou are mostly of this sort. It would seem unprofitable to list them, when I recognize that his views are at least as worthy of respect as mine. I have noted a very few misprints, and also a couple of lapses of the pen, among which I would count the omission of *y* from

the list of sounds which, intervening between *r* (*r*) or *ʃ* and *n*, do not prevent change to *ŋ* (64). Also, in §175 (line 2) read '358 n.' for '357'; in §232 (183, line 8) 'ultérieur' for 'antérieur'; in §250 (line 5) '396' for '397'; in §264 (line 9) '*uw*' for '*uy*'; in §286 (234, line 8) 'Ac.' for 'Ag.'; in §308, note 4 (line 5) 'moyen' for 'actif'; in §363 note 1 the ppp. *grbhūda* is noted, but not *grhūda* (post-RV, at least in compounds).

Slightly less trivial are the following. §214: Does the author mean that the adj. *gopayātya* is based on the 3 sg. pres. **gopayati*? If so, the suggestion seems unlikely, and not easily reconcilable with §365. §334 (277, lines 5-6): That radical medial *a* in the 1 sg. perfect remains short 'par souci de différenciation' (as against 3 sg. with *ā*) seems too teleological, whatever the true explanation may be (the laryngealists propose a historical solution). §399: I doubt the occurrence of enclitic *me* at the beginning of a sentence. On RV 5.14.5 *vētu me śṛṇvad dhavam* see Ludwig's note; *me ... dhavam* goes with *vētu* as well as with *śṛṇvad* (note the accent of this). §457: Cases are cited of 'concatenation' in which a single word is repeated in adjoining stanzas. But more striking is the frequent concatenation by phrases, as in RV 1.32.1c *dhann dhīm (dñv apās tatarda)*, 2a *dhann dhīm (pārvate śisriyāṇām)*; ib. 3d (*dhann enām*) *prathamajām dhīnām*, 4a (*ydd indrāhan*) *prathamajām dhīnām*.

Le dialecte de Dushmani: Description de l'un des parlers de l'Albanie du Nord.

By WACŁAW CIMOCHOWSKI. (Poznańskie Towarzystwo Przyjaciół Nauk, Wydział filologiczno-filozoficzny: Prace Komisji Filologicznej, Vol. 14, No. 1.) Pp. 233. Poznań: Nakładem Poznańskiego Towarzystwa Przyjaciół Nauk, 1951.

Reviewed by ERIC P. HAMP, *University of Chicago*

In a lamentably neglected field such as Albanian almost any happening is an event of major importance; amongst such events, this book, despite several shortcomings, is an occurrence of the first magnitude. It is the first reasonably full-length description (minus lexicon) of an individual Albanian dialect, so far as I know, ever to be done; all previous studies are either fragmentary, amateurish, unsystematically compendious, or at best based on an assumed knowledge of the broad lines of Albanian structure. As an indication of the poverty of our information in this field, I list all unsuperseded scientific dialect treatments of any considerable length that I know of:— H. Pedersen, *Albanesische Texte mit Glossar*; Leipzig, 1895: southern Tosk dialect of Camërija, with brief grammatical introduction. N. Jokl, *Vorläufiger Bericht über die im Auftrag der Balkan-Kommission d. Kais. Akad. d. Wiss. in Wien durchgeführten nordostgegischen Dialektstudien*; Wien, 1915: dialect of Vrapčista. G. S. Lowman, *The phonetics of Albanian*, Lg. 8.271 ff. (1932): dialect of Scutari. C. Tagliavini, *L'Albanese di Dalmazia*; Firenze, 1937: dialect of Borgo Erizzo, near Zara, the only isolated Geg dialect outside Albania; besides about 25 pages of grammatical comment, contains a valuable 235-page *lessico etimologico*, the most up-to-date such tool for Albanian. C. Tagliavini, *Le parlate albanesi di tipo Ghego orientale (Dardania e Macedonia nord-occidentale), Le terre al-*

banesi redente; Roma, 1942: monograph discussing chief distinguishing features. Π. Α. Φουρική, 'Η ἐν Ἀττικῇ ἑλληναλβανικὴ διάλεκτος, 'Αθηνά, Vol. 44-5; Athens, 1932-3: dialect of isolated Tosk colonies in Attica. M. Lambertz, *Italoalbanische Dialektstudien*, *Kuhns Zeitschrift* 51.259-90 (1923), 52.43-90 (1924), 53.66-79 and 282-304 (1925): dialects of six clustered Tosk villages in Molise; comparative grammatical comment, texts, glossary. M. Lambertz, *Albanische Mundarten in Italien*, *Indogermanisches Jahrbuch* 2.1-30 (1914-5): brief general treatment, drawn mainly from the dialects treated in the last item and from scattered written sources. M. Lambertz, *Die Mundarten der albanischen Sprache und ihre Erforschung*, *Leipziger Vierteljahrsschrift für Südosteuropa* 1944. Jaberg-Jud, *Sprach- und Sachatlas Italiens und der Südschweiz*: point 751, Acquafredda, is a Tosk village of Calabria, one of a group of seven located near Castrovillari. K. Kristoforidhi, *Λεξικὸν τῆς Ἀλβανικῆς γλώσσης*; Athens, 1904: old, and in nonstandard Greek orthography, but unreplaced as a reliable source for citations of provenience. — To these may be added:— St. Mladenov, *Die Albaner und das Albanische in Nordmakedonien und Altserbien*, *Balkan Archiv* 1.43-70 (1925). A. Dozon, *Manuel de la langue chkepe ou albanaise*; Paris, 1879: a 'general' grammar based on the Tosk dialect of Përmet. G. Weigand, *Albanesische Grammatik im südgegischen Dialekt (Durazzo, Elbassan, Tirana)*; Leipzig, 1913. G. Pitré, *Fiabe, novelle e racconti popolari siciliane* IV; Palermo, 1875: Tosk dialects of the Sicilian villages of Piana dei Greci (now Piana degli Albanesi) and Palazzo Adriano. In point of grammatical completeness the present study outranks all its predecessors.

Although, as will perhaps become clear by implication in the specific discussion that follows, this description is outmoded in methodology and in form of presentation, and although in certain respects, particular as well as general, it falls far short of the standard demanded by up-to-date linguistic technique, two considerations must mitigate the harsh judgment that more favorable circumstances would otherwise impose. First and foremost, in view of the dearth of material on Albanian, we must be grateful for such a rich and complete picture of a single dialect; no criticism must detract from this fact. Secondly, the material was gathered in 1937 and the analysis carried out during 1938 and 1939, publication having been prevented by the outbreak of the war. We must therefore not look for the results of the great strides made in linguistic science during the past fifteen years. It would be out of place to level empty criticism at many aspects. Let us rather express the hope that our Polish colleagues may once again attain the place of eminence in linguistics once held by their brilliant forefathers.

1. The Introduction (1-5) begins:

Le présent ouvrage contient la description du dialecte albanais du nord, dans la région montagneuse de Dushmani. Ce territoire se trouve à une distance d'environ 30 km., en ligne droite, à l'est de Scutari, avec un écart insensible vers le Nord et il s'étend à droite de la rivière Drin dans son cours inférieur. Ce territoire n'est pas grand, mais il y a lieu d'en distinguer certaines parties composantes, notamment: Malagjia, où se trouve la grotte dans laquelle Fr. Nopcsa fit en son temps des découvertes archéologiques, et Klllogjën. Je mentionne ces deux parties composantes de Dushmani, car j'ai réussi à constater certaines petites différences linguistiques dans leurs territoires. La densité de la population y est petite, selon les données sta-

tistiques de l'année 1936 elle se montait en moyenne à moins de 13 personnes par kilomètre carré. Les montagnards habitant Dushmani sont de religion catholique-romaine.

According to Cimochowski's account, he seems to have been fortunate in having secured an extremely amiable, intelligent, and trustworthy informant, of whom he made very good use, in the town of Scutari, over a period of almost eight months, controlling his data by a two weeks' visit to Dushmani. The author spent an enormous amount of time with his informant, and all cited material comes from that old mountaineer, named Gjok(ë) Pjetri(t). It is to be hoped that at some future time, not too far off, Cimochowski may publish a substantial selection from the large number of texts that he must have collected.

Regarding the setting up of form-classes, Cimochowski says (3) in commendable fashion: 'Cette classification est basée sur des coïncidences de flexion de nature purement extérieure et formelle, il ne faut jamais chercher de base pour cette classification dans la grammaire historique. Une classification en groupes d'après la grammaire historique serait impossible dans beaucoup de cas et absolument inutile pour la description du dialecte.' Yet, after pointing out that his study is descriptive in nature, he remarks (4): 'Si dans certains cas peu nombreux je me suis permis quelques digressions plus étendues concernant l'histoire de la langue, je crois qu'elles étaient indispensables.'

In the chapter entitled *Phonétique*, no distinction is made in any systematic fashion between phonetic and phonemic phenomena. Though in some cases the author is at pains to describe carefully the phonetic value of certain sounds, his general level of phonetic performance is disappointing; e.g. (6): 'Dans le dialecte de Dushmani ... on articule *a* en gardant la position normale de la langue dans la cavité de la bouche ... *i* est une voyelle brève normale de ce type ... *e*, dans le dialecte de Dushmani, est une voyelle brève, ouverte, mais pas dans le même degré que *p*. ex. chez les habitants catholiques de la ville de Scutari.' Any linguist should of course know that there are no universal standards for 'normal' and that the inhabitants of Scutari (what, incidentally, about the non-Catholics?) are not the most accessible and available yardstick for simple phonetic facts. Though the word 'phonème' occurs (e.g. 23, §31), the author shows no sign of having grasped its significance in the senses in which it had already been used for nearly a decade in this country, in Prague, and in Scandinavia at the time when he began his study.

I proceed now to a brief restatement of the phonology.

2.1. Vowel phonemes. Cimochowski lists and illustrates fifty-one vocalic phones: 21 'voyelles buccales' (7 short, 8 long, 6 half-long), 7 'voyelles nasales', 13 'diphthongues avec *i* et *y* à la suite d'éléments normale' (i.e. with the semivowel following: 7 short and 6 long), 7 'diphthongues avec *i* et *y* à la suite d'éléments inverse' (i.e. with the semivowel preceding: 3 short and 4 long), and 3 'liaisons diphthongales' (*ie*, *ue*, *üe*). If we replace Cimochowski's symbols by letters that will better accommodate typographic convenience and existing conventions of official Albanian orthography, the following nuclear syllabics can easily be shown to be minimally contrastive:

A. /a i o y æ ø/; B. /e u/.

The last two members of class A are rare in occurrence; the last three of that class represent Cimochowski's *û*, *ä*, and *ö* respectively. Class B is defined as comprising syllabics that may also occur as unstressed members of stressed nuclei; they could equally well be called semivowels. The classification just given orders the phonemes on the basis of distribution and frequency. On an articulatory basis the same phonemes may be grouped as follows:

	FRONT			
	<hr/>			
HIGH	i	y	u	
MID	e	ø	o	
LOW	æ			a
		<hr/>		
		ROUNDED		

2.2. Prosodic phonemes. By positing two more phonemes, /*ẽ*/ and nasalization, we can accommodate all the remaining elements in Cimochowski's list of fifty-one.

2.2.1. The nasal vowels, so called, can clearly be stated as /V/ plus nasalization, which has the allophonic properties of simultaneous nasality and lengthening. The convenience in separating the nasal component from the vowel, and in not treating the complex as a unit, is confirmed on the phonological level by the resultant class, here set up, of prosodic units that can be manipulated as a single set. On the morphophonemic level, in view of Cimochowski's statement (8) that vowels become nasal before nasal consonants, and (156 fn. 1) that in certain verb classes 'il ne faut pas oublier que la nasalité de la voyelle est très souvent facultative', we find further confirmation of the separation of this prosodic element, whereby such replacements may now be stated merely as combination phenomena parallel to other such combinative alternations.

2.2.2. Cimochowski's length. There is no question but that in this dialect phonetically long vowels contrast distinctively with short; the phonemic interpretation of this length is, however, a matter for discussion. In final position stressed vowels occur only in the form that Cimochowski writes long; in the absence of more precise phonetic detail these may be considered allophones of simple vowels in the sequence /*Ŵ* #/. We therefore find that phonetically long and short vowels are in contrast only in position before consonant /C/.

In a poorly organized section (10-2) entitled *Le traitement de la voyelle sourde dans le dialecte de Dushmani et dans la partie nord de la région dialectale guegue en général*, Cimochowski presents (mixing comparative data with descriptive) certain of the phenomena that in dialects outside northern Albania match a minimally accented short central vowel, which the standard orthography writes *ẽ*. These phenomena include length in the vowel preceding the point under discussion, vowels of varying timbres which correlate with the neighboring C, and transition from voiced to voiceless C—the last feature running counter to the otherwise prevailing assimilated cluster pattern of the dialect. Though the author mentions nothing about the presence of a very short vocoid in such transitions from voiced to voiceless, I should myself very much like to see more detailed evidence on this point. In other Albanian dialects of my acquaintance,

what at first appear to be sundry short, vague transitional vowels amid great varieties of consonants in close succession, later turn out to be classifiable as two distinct sorts of phenomena. One class comprises very short but distinctive vowels which, depending on the dialect in question, may be associated phonemically with either juncture / # / or some vowel (say /ë/); the other class comprises open allophonic transitions between clustered consonants. It would be interesting to know whether such a situation applies in Dushmani, but we need more phonetic data than Cimochoowski provides. It is also not quite clear from his description whether the vocoids that correlate with neighboring consonants, alluded to above, are of such phonetic quality as to be dissociable from the vowel phonemes to which he likens them. It is further not even quite clear from his phraseology whether this 'voyelle brève, fort réduite', which this section discusses, does in fact subsist in any position, in Cimochoowski's view, or whether the vowel mentioned is supposed to exist only in other dialects to which comparative allusion is made. All the phenomena just referred to, on which doubts have been raised and for the solution of which more extensive and intensive phonetic data would be required, will for the sake of brevity be referred to in the ensuing statements as 'possible short vowels'. Finally—and on this point I have also received verbal corroboration from Leonard Newmark of Indiana University—in some dialects the allophone of *ë* in prejunctural position is simply consonant-explosion, which contrasts distinctively with unreleased implosion; it would be interesting to know whether similar phenomena apply to northern Geg.

We are now in a position to formulate a statement for our phoneme /ë/; it has at least the following four allophones: vocalic length, in the sequence /VCë/; transition phenomena from voicing to voicelessness, in the sequence /CëC/; possible short vowels, as above discussed, in the sequence /CëC/; possible consonant-release phenomena, as above mentioned, in the sequence /Cë #/. Thus, *löp* 'cow' is /löpë/; *šüm* 'much' is /šümë/; *vāḡzs* 'of the girl' is /vāḡzës/. In contrast with *čime* 'cheveu' and *hiḡe* 'ombre', we may interpret *e čime* 'coit' and *pīḡe* 'boisson' as /ečímëe/, /pījëe/. Thus interpreted, though in many respects our /ë/ corresponds to the vowel phoneme *ē* of other dialects, its distribution differs significantly, since in those dialects it does not occur immediately before /V/.

2.2.3. Cimochoowski's half-length. According to the author, this feature is restricted almost entirely to two morphological classes: verbal infinitives and the definite form of certain nouns. For a phonological feature, such a morphologically defined distribution seems unlikely; it is at any rate analytically unsatisfying, and should be rechecked. It is further stated that the feature occurs also in some isolated words, but the only example adduced is *termëk* 'earthquake'. Since the word is attested in other dialects as *termët* (a very old loan from Latin), is it not possible that this has been reinterpreted by folk-etymology as an infinitive, here in Turkish form and perhaps connected with *tepremëk* 'stir, move'? If so, it falls under the above rubric.

In any event, we may tentatively interpret half-length as /Vë/. Thus *me vēḡ* 's'habiller' is /mevëḡs/, etc.

2.3. Syllabic combinations and complex nuclei. We may now reinterpret Cimochoowski's remaining elements as follows. *Vḡ* is simply /Vj/, since a phoneme

/j/ is in any case required on other grounds; the symbol is selected on the basis of the standard orthography. $V\check{u}$ is / $\check{V}u$ /; the non-occurrence of $*\check{V}u$ is an interesting confirmation of the lack of evidence in other positions for $*w$, as contrasted with /j/. $\check{V}\check{i}$ is / $\check{V}j\check{e}$ /. $\check{i}V$ is /jV/. $\check{i}\check{V}C(C)$ is /j $\check{V}C(C)\check{e}$ /. The sole case with $\check{u}\check{V}$, *kuál* 'horses', is /kuál \check{e} /. In all of the above, /j/ is of course merely a particular instance of /C/. Finally, the occurring instances of $\check{V}e$ are / $\check{e} \check{u}e \check{y}e$ /.

2.4. Consonant phonemes. Cimochowski's consonants are here briefly recapitulated with some comment.

/p/ is lenis in final position, fortis elsewhere. /b/ is lenis, but is said to be more fortis before a front vowel. /t/ is described as fortis in all positions. /d/ is stated to be voiced, but nothing more. /k/ is lenis in final position, fortis elsewhere. /g/ is stated to be voiced and unaspirate, but nothing more. In final position, voiced are replaced morphophonemically by voiceless (lenis), but in this respect the author fails to differentiate the phonological from the morphophonemic level in his statement.

/j/ is a non-syllabic vocoid. Cimochowski also reports a supposedly distinct 'spirante j', which is stated to occur in only two words *zoi* -t 'birds' and *i kolájt* 'easy'; this would appear to be a special allophone or free variant of / \check{z} / before /t/ (or before / $\check{e}t$ /?). /l/ has two major allophones: [l] before /a/ and back vowels; [l'] before /i y/. /l/ is strongly velarized and with incomplete occlusion between tongue-tip and upper teeth; in some words, interchanges occur between dialects between this phoneme and / $\check{\delta}$ /. /r/ is short, weakly rolled, retroflex. / \check{r} / is a long fortis velarized rolled sound.

/m/ labial nasal; the author states that there is one lone form with palatalized *m̃*, *me m̃el* 'to milk'; but with such a distribution it is hard to believe that this is not a free variant of, say, /memj $\check{e}l$ /. /n/ dental nasal. / \check{n} / palatal nasal.

/v/ is stated to involve two simultaneous spirant occlusions: lower lip against upper teeth and upper lip in front of lower lip. /v/ has a faulty distribution; it does not occur prejuncturally. Since only [u] occurs in that position, it is simpler on grounds of phonetic similarity, as well as on the basis of morphophonemic distributions, to interpret the latter as /u/, which regularly replaces /v/ in this position. /s/ and /z/ require little comment. Prejuncturally /s/ always replaces /z/; *e z $\check{e}z$* 'black (fem.)' and *unáz* 'ring' do not alter this statement, since they are phonemically /ez $\check{e}z\check{e}$ /, /unáz \check{e} /. / \check{s} / and / \check{z} / raise only one point: Three words (irregular aorist 1 sg.) are said to have a final \check{z} '; it would appear most likely that this is a free variant of / \check{c} /. /p/ and / $\check{\delta}$ / are dental spirants with considerable lateral opening. The prejunctural allophone of / $\check{\delta}$ / is devoiced during the final part of its span. /f/ is labiodental, with the tongue simultaneously in the position it assumes for /p/; this serves to explain the occasional inter-dialect alternations between these two phonemes in various words. /c/ and / \check{z} / are sibilant affricates. / \check{c} / and / \check{z} / are shibilant affricates, and, as the author notes, contrast with the sequence /t \check{s} /. / \check{c} / and / \check{z} / are palatalized shibilant affricates; they correspond to the palatal stops of other Albanian dialects (orthographic *q* and *gj*). Prejuncturally / \check{c} / replaces / \check{z} /. /h/ is a pharyngeal-laryngeal spirant with two major allophones, voiced postjuncturally and voiceless prejuncturally.

Two phones occur which Cimochowski writes *k'* and *g'*, and which like several other interesting features he discusses only in a comparative-historical setting; these occur only before /j/, and appear to be allophones of /k g/ in that position.

2.4.1. Summary of consonants. The consonant phonemes may be tabulated structurally on a combined basis of phonemic and morphophonemic distribution; though their phonetic properties interestingly correlate with this arrangement to some degree, the labels applied to the classes are merely a matter of convenience.

	STOPS			AFFRICATES			S(H)IBILANTS	
VOICELESS	p	t	k	c	č	ć	s	š
VOICED	b	d	g	3	ž	ž	z	ž
NASAL	m	n				ñ		
LOWER-CLUSTERABLE SPIRANTS				v	ǰ	ɸ	f	h
HIGH-CLUSTERABLE SONANTS				r	ř	l	l	j

2.4.2. Consonant combinations. We cannot attempt here to deal with the rich array of occurring clusters in any systematic fashion. A few points, however, demand comment because of the way in which the author's quasi-comparative presentation occludes their descriptive status.

Although in general /pj bj fj/ occur, occasionally the more archaic /pl bl fl/ are found in their stead. This looks like dialect mixture.

Complexes that correspond to nasal plus (voiced) stop in other dialects (phonemically a unitary prenasalized stop in one Calabrian-Albanian dialect of my acquaintance, Vaccarizzo Albanese, and a unitary distinctive long emphatic nasal in some of the northern Geg dialects) are treated here as follows, as far as I can make out.

A sequence [ŋ²] occurs, with a very brief weak stop. This may be interpreted as the phonemic sequence /ng/.

Though most instances of *mb* in other dialects are matched by a simple /m/ in Dushmani, according to the author's account, there are some occurrences which he writes *m^b*, without making their status clear and without stating conditions for their occurrence; whatever lies behind them—and they look suspiciously like loans or dialect mixture—they are probably best interpreted as /mb/. Cimochowski further remarks that forms occurring in other dialects with *mbš-* have *pš-* in Dushmani. Apart from the fact that this point has no place in a descriptive phonology, Cimochowski's proposed phonological explanation does not accord with the facts, since he cites other forms that do not behave this way. Instead, it would seem as though forms such as the cognate of Tosk *mbështjellur* 'wrap up' were misanalyzed as belonging to the common verbal pattern with simple nasal prefix, and a falsely decomposed simplex extracted as /pštjéél/.

Though practically all cases of *nd* in other dialects are matched by simple /n/ in Dushmani, Cimochowski claims that in two forms (*me ññe* 'feel, hear', *tañe* 'thy acc. fem. sg.') one finds a long emphatic ñ, as in Scutari. If this is to be taken at face value, the span can no doubt be reinterpreted as the special allophonic form of /nd/; more data on this would be desirable.

All cases of *ngj* in other dialects are matched by simple /*nj*/ in Dushmani, except for those that go back to common Albanian *njl* (preserved in some of the isolated dialects of Greece and Italy), which is represented by /*nj*/ . This last cluster, incidentally, is one of the best reasons for not reinterpreting /*nj*/ as /*nj*/.

2.5. Stress phonemes. Although Cimochoowski overtly mentions (27) only one stress, forms such as *grúesmîr* 'good woman's', *grávetmîra* 'good women's' (73) make it obvious that there are three significant degrees of stress, counting minimal, which we may write /' ^ ~/. The above forms may thus be rewritten /*grúes-mîrê* #/, /*grávêët-mîrêä* #/; for convenience, minimal stress may be left unmarked, though this of course does not mean that it does not exist as a significant element.

2.6. Summary of phonemes. We have set up 8 vowels and semivowels, 29 consonants, and 5 prosodic and stress phonemes; the last, which are here grouped together, all structure together as a larger set, since the occurrence of the prosodic units bears a relationship to that of non-minimal stress. Though Cimochoowski takes no account of such things, it is likely that there are two or three intonation levels of phonemic significance, and at least three junctures and terminal contours; let us allow for half a dozen phonemes of pitch, juncture, and contour, however they may structure with the other phonemes.

The roster of Dushmani phonemes then stands at about fifty. It is a pity that in a monograph the size of Cimochoowski's we cannot find collected data adequate for a statement of the occurring sequences of all phonemes up to some manageable number, say three.

3. Noun morphology. The Albanian noun is complicated, and Dushmani embodies most of the interesting complexities. Cimochoowski gives a very detailed and (on the whole) careful account; after discussing (28-36) gender and number, and illustrating these in detail with types of noun-membership, he presents (37-70) the nominal inflexion, with a rich array of listings of members of the numerous small classes. Cimochoowski's general arrangement is quite clear: the category of definiteness, illustrated paradigmatically by the main inflexional types; the so-called 'article', or, as I should prefer without wishing to quibble, the prefix of concord; the separate noun types, by gender and class of plural formation, with paradigms and lists of members for each. By discussing and illustrating the manner of linking two nouns together (i.e. constructing a nominal phrase), the author leads conveniently into his succeeding chapter on the adjective. In reality, though there will not be space below to deal with this aspect, in Albanian the adjective is merely a subclass of the substantive, and its description belongs primarily in the section on nominal derivation and in that on nominal phrase construction. It is probably simpler, too, in Albanian to make the basic unit of description the nominal phrase, whether or not a juncture intervenes, as it often does not, since the features of inflexion and concord occur throughout this unit and on both ends of it, and are most readily described in one ranked set of statements. Thus the single noun (and substantivized adjective) is merely the special limiting case of the nominal phrase.

Though one cannot deny the complexity of Albanian noun structure, with its numerous small classes and morphophonemic interchanges, and though Cimo-

chowski offers a rich and useful collection and classification of material, his presentation is needlessly uneconomical, with forty-five separate noun paradigms. It is usually possible to find most things one is apt to look for, but a clear picture of how the machine works fails to emerge in its broad lines. An attempt is made below to restate the noun, without, however, taking the space to define the categories set up. So as not to interfere annoyingly with the already compressed presentation, specific paragraph references to Cimochofski are dispensed with.

3.1. Morphophonemics. The following symbols, which are used as covers for phonemic interchanges, and the following phonemic relationships must be established in advance. K = /k g/ are replaced by /č ž/; I = /i/ is added before /t š/; E = /e/, when I does not vocalize and when /j/ does not follow in Class II, in other circumstances being equal to zero; J = /j/ replaces /n l/, and in other circumstances is added as a suffix; 1! = stem vowel /a/ is unlauted to (replaced by) /e/; 2! = any stem vowel is unlauted to /i/; 3! = stem vowel /e/ is unlauted to /a/ (applies only to /pés/ 'bag' and /rēp/ 'circle').

Any geminate C produced by suffixation simplifies to C. Except in certain cases, /-ë/ is lost before suffixed V. After -V̇ a suffixed C is replaced by /Cë/. After /k g a e i/ a suffixed /i/ is replaced by /u/.

3.2. The nominal form-class. The nominal phrase may contain only one nucleus, and the case and definitizing suffixes are limited to the nucleus, except for those labeled 'marginal'.¹ A noun is a base (plus suffixes) that may occupy such a nucleus. A nominal phrase nucleus must bear one and only one case suffix; any bases following the nucleus may, but need not, bear a marginal suffix. With the exception of certain specific morphemes (e.g. /žip-/ 'all'), the first base in a nominal phrase is the locus of the nucleus. Noun bases that may occur after the nucleus without a marginal suffix, but always with a prefix of concord, may be called adjectives; these bases have the further peculiarity of occurring in all possible gender formations and concords. With certain classes of exceptions, adjectives and bases with marginal suffixes regularly take a prefix of concord, the concord being governed by the preceding number-gender-case-definiteness suffixes. A nucleus, or other noun-base position, may also be occupied by a pronoun, which is another rather complicated and irregular, but numerically small form-class; a pronoun may also occur before the nucleus. In other words, pronouns form a different form-class from nouns, but may enter into nominal phrases. A certain number of particles (adverbs) may intervene in, and form part of, nominal phrases. Adjectives, and thus nouns, may be regularly derived from verbal bases. Every nominal phrase probably contains a superfix of stresses; the favorite superfix, from what little we can glean from the author's illustrations, would seem to be /[^].../.

Some nouns take nominative and accusative formations which, though sin-

¹ I take this term from Leonard Newmark, who introduced it in his interesting paper, *An Albanian case system*, presented at the meeting of the Linguistic Society in Bloomington, Indiana, 8 August 1953. The Dushmani system differs in some respects (and in others has been differently interpreted) from that described by Newmark.

gular in meaning, are plural in form (= 'neuters'). As an irregular submember of this group, the Class II noun /kre-/ 'head' has the plural-form nom./acc. /krýe/.

The singular oblique cases of /-at-/ 'father' have thoroughly irregular port-manteau forms. Otherwise, all noun bases can be managed readily with the morphophonemics stated and with a minimum number of allomorphs.

3.3. Order of affixes. The fullest possible form of a noun is

$$P_2P_1(P)NN(SS \dots)S_1S_2S_3,$$

where P stands for prefix, N stands for noun base, and S for suffix. The symbols in parenthesis allow for derivational affixes on noun bases (or bases that have already been nominalized). NN allows for compound nouns, which occur. Occupants of the other affix positions, which are the features that define the noun form-class, are listed:

P_2 : prefix of concord, definite or indefinite in form, with the concord governed by preceding suffixes.

P_1 : prefix of concord, indefinite (except after particle /ma/, with which it occurs definite and has superlative meaning) in form, with concord governed by following suffixes. This prefix occurs with (a) certain masculines and feminines of relationship, (b) all 'neuters' except /kre-/, and (c) substantivized adjectives.

S_1 : plural, definite singular, marginal singular, marginal ablative singular.

S_2 : nominative, accusative, marginal plural.

S_3 : definite plural, marginal ablative plural.

All affixes within a position as stated are mutually exclusive. With the proviso stated earlier that only one case suffix may occur, all positions may be present in all combinations.

3.4. Suffixes. Two classes of nouns are distinguishable on the basis of the suffixes they select; this is the state of affairs, rather than the reverse, namely, that the terminations are selected on the basis of gender, as Cimochowski has it (37).

Class I comprises masculines, certain subgroups of which, however, are feminine in the plural, as is noted below with the relevant suffixes. Class II comprises chiefly feminines and a small number of masculines. As will be seen, the class as here set up refers to suffix-selection, while gender refers to concord within the total nominal phrase.

The suffix forms are shown in the table on the next page.²

The marginal singular, the only case that behaves asymmetrically with respect to definiteness, is treated as above for the following reasons: (a) definiteness finds no opposition in the Class I nouns; (b) the above statement saves positing complicated allomorphs for the definite singular; (c) it saves obtruding complexities into the clear-cut plural structure; (d) it relegates a minor and vestigial structural opposition to its appropriate status—a minor place in the table.

3.5. Plural affix allomorphs. The numerous allomorphs of the plural affix are responsible for setting up many small classes of nouns whose membership simply

² Superior letters in the table have the following meanings: a = before nominative; b = before accusative; c = indefinite; d = definite.

	CLASS I:	CLASS II:	
		Bases in -V	Other Bases
plural (S ₁)	----- See below. -----		
def. sg. (S ₁)	/-i/	/-ja/ ^a	/-a/ ^a
		/-o/ ^b	/-e/ ^b
def pl. (S ₂)	-----/-t/-----		
Cases:			
nom. (S ₂)	-----/-o/-----		
acc. (S ₂)	-----after def. sg. /-n/-----		
	-----elsewhere.../-o/-----		
marg. sg. (S ₁)	/-it/ ~ /-t/	/-je/ ^c	/-o, -s/ ^c
	(<i>vlā</i> 'brother')	/-s/ ^d	/-s/ ^d
marg. pl. (S ₂)	-----1!-EnI class /-e/-----		
	-----elsewhere.../-ve/-----		
marg. abl. sg. (S ₁)	/-it/	/-jet/	/-et/
marg. abl. pl. (S ₂)	-----/-š/-----		

has to be listed; they contribute greatly to making the Albanian noun the complicated picture it is when described in extenso. The forms that occur are simply listed below, together with the paragraph numbers where Cimochowski lists members of the class in question; the spread in numbers will give a notion of how uneconomical his classification is—a classification partly based on what phoneme the singular happens to end with.

It will be seen that a single plural affix may consist of a complex of up to three affixal allomorphs. The members of a class range in number from one up to hundreds; though space does not permit the inclusion of membership totals for each class, suffice it to say that the elimination of classes of less than five would not materially change the complexity of the structure as stated.

Twenty-four allomorphs and allomorph combinations occur; the latter are broken up in the list below by hyphens. An asterisk marks allomorphs that apply to nouns of both Class I and Class II; a roman II, those that apply to nouns of Class II only; a dagger, to 'neuters'. Allomorphs that apply to nouns of Class I only are unmarked. References are to Cimochowski's paragraphs.

/-o/* (§§47, 59b, 59c, 60, 61, 65-8); /1!/* (§§57, 58, 59a, 70); /3!/ (§57); /VV/: /ue/ > /a/, /o/ > /ue/, /e/ > /ye/ II (§70), /a/ > /ua/ (§60); /-ē/ (§§48, 53); /K/ (§§49b, 85); /J/* (§§52, 58, 64a, 85); /-a/*† (§§50, 58, 59a, 62, 64b, 69, 71, 85); /-na/* (§§54, 55, 63, 65); /-e/ (§51). The last two form feminine plurals regardless of the gender of the base in the singular.

/J-a/ II (§§64b, 67, 85); /K-I/ (§49a); /1!-K/ (§49b); /J-e/ (§51d); /-jJ-e/ (§51e); /K-e/ (§51f_α); /1!-K-e/ (§51f_β); /2!-K-e/ (§51f_β); /1!-na/ (§§54, 55); /2!-na/ (§56); /1!-EnI/ (§56); /1!-mē/ (§57); /-z-EnI/ (§58); /-z-I/ (§60).

The same elements can be arranged more economically in a diagram. Reading from left to right without crossing the horizontal lines, all the elements occur singly except those in parentheses, and all combinations occur (except a few

involving parenthetical elements). A few combinations are here more easily treated as units.

0,	VV,	3l,	ě,	KI
zI			(EnI)	
			(mē)	
11, (21)			na	
		K		
(i)	J		e	
			a	

3.6. Concord prefix allomorphs. The following forms occur for the stated concords: /e-/: acc. def.; nom. pl. def.; nom. sg. fem. /i/: nom. sg. masc. /s-/
(~ /t-/ with superlative): marg. sg. def. fem.; marg.-abl. indef. fem. /t-, s-/:
marg. sg. indef. fem.; marg.-abl. indef. masc.; marg.-abl. sg. def. fem. /t-/: bal-
ance. For 'neuters': in the nom., the sg. equals the pl.; the balance of the sg.
equals the masc.; all the pl. equals the fem.

3.7. Illustrative nominal phrases. One example will have to suffice; we take
the utterance given by Cimochowski (84) as *grúe mą e mīra e t tǎnavet* 'la femme
la meilleure de toutes'. Using all available data, this is transcribed as /grúe #
mǎ # ěmīrěǎ # ět(#)tǎnǎvēt # /. Morphologically, this comprises two nom-
inal phrases, consisting of /grúe # /, and the balance; perhaps *mīra* should also
have a primary stress. The first simple phrase comprises the nucleus /grue/
'woman', plus S₂ /-0/ (nom.).

The balance may be symbolized as X # P₁N_aS₁S₂ # P₂P₁N₅S₅S₅. X is a par-
ticle meaning 'more'. P₁ is definite (with superlative meaning) and in concord
with following suffixes. N_a /mirě/ 'good' is an adjective (here nominalized),
and is the nucleus. S₁ /-a/ (def. sg. allomorph before nom. after non-vowel).
S₂ /0/ (nom.). The whole is Class II, fem. P₂ is /e-/ (nom. def. sg. fem.), in con-
cord with the preceding forms. P₁ /t-/ (marg. def. pl. fem.) is in concord with the
following morphemes, though, as it happens, with /tǎn-/, which always has P₁,
the form is fixed as /t-/ for all concords. /tǎn-/ is a special type of noun which,
like /ǝip-/, may stand before the nucleus, though in the present instance it
behaves like any N. S₁ /-a/ (plural, perhaps the most common type) inflects
the base as a Class II fem; this changeability of base and gender, as well as other
behaviour, parallels /ǝip-/ and the numerals (which also may precede the nu-
cleus). S₂ /-ve/ (marg. pl.), case suffix that may appear on a non-nucleus. S₅
/-t/ (def. pl.).

4. Balance of morphology. There is not space here to deal in similar fashion
with the rich and interesting verb morphology, which like that of all Albanian
dialects comprises about as many forms as in a Romance language, with pre-

fixes (including fused pronoun objects), suffixes, and (to a more elaborate extent than with the noun) internal vowel and consonant alternation, plus outright suppletion. For all of these the author gives very full paradigms in his section on the verb (109-95). A brief section on syntax closes the book. No systematic treatment is given to derivation processes.

5. Comparative. Many comparative and lexical points, of great interest to students of Albanian, are raised by this book. I shall take up only one here, a point raised in a footnote (108), where derivational prefixes on verb bases are under discussion. Regarding /me-ñíët/ 'encoller', /me-u-ñíët/ 'monter', beside /me-u-žíët/, Cimochowski remarks that this base has /ž-/ instead of the /gj/ that would be expected from earlier *gl-* (cf. above, §2.4 end and §2.4.2 end), and compares Greek-Albanian *glis* (1 sg.). The situation seems to be more complicated than that; it will be of advantage to inspect cross-dialectally a group of several words. From Italo-Albanian I use only forms that I have collected personally from native speakers.

	'dip, infect'	'resemble'	'stick'	'satisfy'
Vaccarizzo	ngí	gít, gásen	ngít-ín	(ngós)
Barile		gl'ft-ən	ngft-əm	
Piana		glft-en		
Dushmani	ñýe, žýe		ñít	ngí
Rossi dict. (N. Geg)	zdžüej		ñit	ng'i
Borgo Erizzo		ngjet		ngiñ
cited in Meyer, <i>Et. Wb.</i>				'Sic.' ngi'ín

On the basis of these we reconstruct the following shapes for the four words in the Albanian of about 1400 (with T for /t/ ~ /s/ and V for vocalic interchange): */n-gý-/ , */n-glV-T-/ , */n-giT-/ , */n-gli-n-/ . In Early Albanian we find for 'resemble' *gla* (Buzuk) and *ga* (Bogdan). In other words, from the evidence assembled where such phenomena show up distinctively (i.e. Barile and North Geg, including Dushmani) it would appear that 'stick' did not have initial **gl-*. Meyer (*Etym. Wb.* 309), under the entry *ngít*, says 'für *gl'it*'; but he cites no evidence and merely adduces alleged cognates.

Because 'stick' in the mediopassive means 'climb', it is thinkable that two old bases have been conflated, perhaps at a time before the separation of the dialects as we now have them. Perhaps the earlier situation was something like */n-giT-/ 'stick', but */n-gliT-/ 'climb'; */n-giT-/ could be related to *qij* 'have sexual intercourse', while */n-gliT-/ could be related to Gk. κλίω and all its numerous relatives.

Irish dialects and Irish-speaking districts: Three lectures. By BRIAN Ó Cúiv. Pp. 94, with folding double map. Dublin: Dublin Institute for Advanced Studies, 1951.

Reviewed by ERIC P. HAMP, *University of Chicago*

Scholars of Irish have come to expect good things of Ó Cúiv, and the present little book is up to standard. In a way, because of its excellence, one might wish

that he had included even more material and made it a fatter book, yet in another way one is grateful that so much has been made compactly and readably accessible to general readers in so small a space.

As the title and subtitle suggest, the book is a collection of three lectures on the extent and shrinkage of the native Irish-speaking area, the so-called *Gaeltacht* (phonemically /géhltəxt/ in the English of Ireland), and on certain noteworthy features of Cork Irish treated in their dialectological setting. The first lecture, entitled *The Gaeltacht—Past and Present*, reviews in twenty-five pages, with admirable clarity and with apt selection of evidential source material, the shrinkage of the Gaeltacht over the past four to five hundred years. This account has an interest for general readers not only as a factual statement of a familiar ethnic-linguistic distribution that has often been the subject of nationalistic distortion and uncritical misstatement, but also as a sample of bilingualism and the competition of linguistic systems in an area where we may hope to know more of the past and to study in greater detail the current development of the total situation than we may reasonably hope for in many parts of the world. It may also be, if we are industrious enough, that we shall have the opportunity here to record and analyze scientifically the grim spectacle of the death of a language; if that day comes, we shall thank Brian Ó Cúív, amongst our foremost creditors, once more for our gains.

The second lecture (33–55), *Irish a Living Language*, is devoted to placing certain dialect features in their diachronic setting as illustrations of linguistic change. Since Ó Cúív, in the happy combination of native speaker and sophisticated linguist, urbane marries hard fact with good sense, the general reader will probably read this lecture with pleasure and profit, even though the content is highly particular at times. Needless to say, the Kelticist will find many an item to interest him and to suggest further studies.

The third lecture, *Some Aspects of Cork Dialects*, gives in seventeen pages a wealth of detailed dialect information, including a fine sketch of the main features of an interesting hitherto undescribed dialect, all of which will be of great use to students of Irish linguistics, though it may not much interest the general reader.

The text concludes with four pages of informative footnotes, an Appendix (77–93) giving county by county, often with tabular breakdowns, a summary of the known extent of Irish in the 19th century, and (94) an Addendum, giving the recently released county census of the Gaeltacht for 1946. An inexpensively printed fold-out double map, showing the density of the Gaeltacht according to the censuses of 1851 and of 1891, closes the little volume, a paper-backed affair with fairly good inkable paper and pleasing typography in conformity with the Institute's usual high standards of clarity and accuracy.

We may now turn to examine and summarize parts of the content. Some quotations will serve to point up the leading facts of the first lecture: 'For the Gaeltacht, which in 1170 was coextensive with our country, has in the course of centuries been gradually reduced until today it has a precarious existence in parts of seven of our thirty-two counties' (8). '... all the evidence available goes to show that' in the 16th century, 'with the exception of a small number in parts of Lein-

ster and in certain urban areas, the people of Ireland were Irish-speaking and Irish-speaking only' (14). 'Lord Deputy Sidney had made a similar suggestion to Elizabeth in 1576, so it seems that Scottish Gaelic speakers were intelligible to the Irish at this period' (15). Speaking of the early 1800's: 'When we consider these estimates we cannot doubt that Irish was spoken by well over two million people, and that at a time when the population of the country was increasing by leaps and bounds. For ... the population of Ireland was estimated in 1788 at just over 4 million souls, whereas the Census of 1841 showed a return of over 8 million' (22). 'According to the Census [of 1851] the total number of Irish speakers was 1,524,286 (or 23 % of the population), of whom 319,602 were Irish-speaking monoglots. The percentages for the provinces were Connacht 51 %, Munster 44 %, Ulster 6.8 %, and Leinster 3.5 %. When we look at the returns for the different counties we get a shock. The highest percentages were Galway 70, Mayo 66, Waterford 63, Kerry 61, Clare 60, Cork 53, Sligo 38, Limerick 37, Tipperary South Riding 31 (with only 4 % in the North Riding), Donegal 29, Roscommon 27. ... In Westmeath, where according to Stokes in 1806 Irish was mostly spoken, there were now only 921 Irish speakers (or 0.8 %). Even allowing for underestimation it is clear that a big change had taken place' (23). 'To sum up: a century ago we would have found Irish speakers in every county in Ireland' (26). 'Irish was still in general use in many towns ... Although 77 persons in every 100 were recorded as being unable to speak Irish, only 5 persons in every hundred were recorded as being unable to speak English ... Far more females were ignorant of English than males' (26). 'By the end of the century the Irish-speaking areas were mainly to be found in Clare, Cork, Kerry, Waterford, Galway, Mayo, Sligo, and Donegal; to a lesser extent in Limerick, Roscommon, and Tipperary; with small patches in Antrim, Armagh, Cavan, Derry, Kilkenny, Leitrim, Louth, Meath, Monaghan, and Tyrone' (27). 'In assessing fully the deterioration we must take into account three other factors—firstly the probable overestimation of 1925 as compared with the underestimation of 1851; secondly that as stated in the Gaeltacht Commission Report the enumeration of 1925 was not concerned as to whether those described as Irish Speakers did in fact use the Irish language as their normal medium of expression; and thirdly that the total number of Irish-speaking monoglots had dropped to 12,000' (29). He concludes his statistics (31-2): 'According to the Census of 1931 there were over 100,000 Gaelic speakers in the Highlands and Islands of Scotland ... Over 900,000 persons, or one-third of the population, are Welsh-speaking, and more important still Welsh is an urban as well as a rural language ... today probably no more than 35,000 persons use Irish as their ordinary medium of speech, and no more than 3,000 are ignorant of English.' It would be empty to add words to such figures.

The second lecture starts with a fine brief statement of the inexorability of linguistic change and of the fact that language is what people speak; this is of course old stuff for readers of this journal, but it must be remembered that in cultures where a major subsystem is losing currency the secondary and tertiary responses may acquire a richer network of semantic correlates than the

forms of the subsystem itself have. In speaking of linguistic age levels, Ó Cuív notes (38) that whereas older speakers in Ballyvourney, West Cork, have a palatalized dental occlusion for their /d' t'/, most younger speakers use an alveolar occlusion, 'no doubt under the influence of English'. I have observed in Waterville, Co. Kerry, a place where only a few stray old speakers are left, that alveolars are regularly substituted in such situations. I further noticed that old Michael Sugrue used a bilabial /f/, while his son-in-law Cornelius Fitz Patrick employed a labiodental.

It is well known that Old and Middle Irish show almost nothing of dialect variation, but Ó Cuív wisely points out that the vast linguistic changes that we see from century to century could scarcely have occurred with absolute uniformity from place to place. At any rate, whatever the result by Middle Irish times, the standard bardic dialect and orthography which was sole possessor of the field from the 13th to the 17th century served to mask almost entirely any dialect variation, however great, in the actual speech. It is only the rarest of 'misspellings' from this period that can be turned to account. Nevertheless, as Ó Cuív points out, the 15th-century annals known as MacCarthaigh's Book, which were written in Co. Cork, do show some Corkisms even at that date. There we find *a-* occasionally written for orthographic initial *ea-*; this correlates well with the known development of Old Irish /e/, which had the allophone [e] ~ [æ] at an early date before a non-palatal consonant, and which in initial position, by passing through the stages [a] and [a], has fallen together with /a/ in this position in present-day Ballyvourney. Several interesting dialectal peculiarities are to be found in 18th-century manuscripts written by Seán Ó Murchadha na Ráithíneach, of Carrignavar in mid-Cork. As Ó Cuív remarks (45), 'in almost every way the Irish pronunciation in mid-Cork two hundred years ago was very similar to that in Ballyvourney today.' And we may safely say that a good many of the features go a long way further back than that.

By inspecting a catechism of 1792, Ó Cuív is able to set forth in a few paragraphs a fascinating little network of similarities and dissimilarities to the modern Irish of Ballyvourney, in West Muskerry, and adjacent regions of Cork and Kerry. Though his modern control for these comparisons is the speech of West Muskerry, he asserts (47) 'from what information I have been able to glean from the last native speaker from Carrignavar I can state that in its last stages the dialect of mid-Cork seems to have differed little from West Muskerry Irish.'

After discussing a set of isoglosses for Munster Irish, Ó Cuív makes a statement (53) which, together with those immediately following, is worth quoting in full: 'This phenomenon of the correspondence of forms between areas far apart is one which is difficult to explain satisfactorily. Another example of it is the use in Clare and in Donegal of *f* forms of the future and conditional similar to those that I described as being current in South-west Cork. Perhaps the most extreme case geographically is the lengthening or diphthongisation of vowels in words like *fiúnn*, *ím*, and *claunn*, in Northern Scotland and South-west Munster.' Now Meillet used to like to refer to correspondences that on first sight had the appearance of being common inheritances, but that on closer inspection

are shown to be so-called independent developments along parallel lines because of common-source 'tendencies'. I think that our understanding of structural phenomena has now deepened to a point where we can attempt a programmatic answer to such riddles in comparative and dialectological study. We now know that languages have not only a phonemic structure, but also a subphonemic, or allophonic, structure. Because the latter is not distinctive ON ONE LEVEL, it does not mean that it is unstructured; it is, if you like, the redundancy structure, the counterpart, on one side, of the primary structure. With the wretched orthography that we find in Old Irish, it is a tricky business to extract the phonemic system; but in the process of doing that, a fair amount of the allophonic structure comes out at the same time.

It is clear that in just about all positions the Old Irish emphatic ('nonlenited') liquids and nasals /L R N ŋ/ and /m/ (no matter whether the latter was structurally an emphatic or a buccal stop) had allophones distinguished by length as well as by point of articulation. It is also clear that allophonic glides occurred regularly between vowels and consonants when the vowel had an articulation different from the simultaneous component that occurred with every consonant. In the case of the consonants in question it is likely that these glides were especially prominent, and since these consonants were in any case distinguished by mode of articulation, their 'extra' resonance in the length was always available for restructuring along with the glide. When, furthermore, in Middle Irish the distinction disappeared between the lip-rounded and the non-palatal components, the free variation in these glides became potentially even greater. It is not hard to see, then, that very little independent change was required in the separate dialects so as to revalue these preexisting configurations.

We thus see that the shared feature was there all the time; it has merely been retained at the expense of some other features and moved onto another structural level in certain of the dialects. Since in the case in point there was a limited number of alternative structural movements open in Middle Irish, we might expect to find amongst all the modern dialects a few agreements, however widely scattered geographically; there should be a statistical way of evaluating such agreements. It is only shared innovations which are not immediately inferable from preexisting nondistinctive structures that determine genetic relationships. (On levels other than the phonological, proof of this point often becomes complicated indeed, but the principle remains the same.

In the last lecture Ó Cuív surveys the dialects of Co. Cork on the basis of some very thorough field work recently done by him, and includes a sketch of the interesting characteristic features of the Irish of Ballymacoda, East Cork (61-8).

His conclusions (70-1) on the subdialects of Co. Cork are that there are three main groups. The coastal region 'extends from Kilbrittain on the coast south-west of Cork city through Carbery westwards into Berehaven'. The central region covers Mid and West Cork, with Carrignavar the most easterly point recorded. 'Finally we have the southeastern region ... This area was cut off from the south and south-western coastal region by Cork Harbour and Cork City, a

factor which probably led to closer contacts with the adjoining county of Waterford which are clearly reflected in the dialect.'

The Irish of West Muskerry, Co. Cork: A phonetic study. By BRIAN Ó CÚFV. Pp. xi, 159, with Corrigenda. Dublin: Dublin Institute for Advanced Studies, 1944.

The Irish of Cois Fhairrge, Co. Galway: A phonetic study. By TOMÁS DE BHALLDRAITHE. Pp. xii, 153. Dublin: Dublin Institute for Advanced Studies, 1945.

The Irish of Ring, Co. Waterford: A phonetic study. By RISTEARD B. BREATHNACH. Pp. xviii, 176. Dublin: Dublin Institute for Advanced Studies, 1947.

Reviewed by ERIC P. HAMP, *University of Chicago*

There are several good reasons for taking this opportunity to review these three books, despite the lapse of time since their publication. The chief reason is that they are eminently good and useful contributions to the world of scholarship on the Irish language and to the growing body of reliable phonological dialect descriptions. The authors and the Dublin Institute for Advanced Studies are to be congratulated on a fine performance all round. It is furthermore gratifying to see native Irishmen with an intimate knowledge of the total milieu turning out such fine specimens of up-to-date methodology. Differing as they do in interesting ways from the better known surrounding European languages, all dialects of Keltic have a certain enhanced intrinsic interest. Finally, these descriptions may be taken to represent the application of a methodology typical of one 'school' of linguists, that associated with the name of Daniel Jones; it is of interest to inspect a few of the more general theoretical implications of these works.

1. Format. The three volumes are exactly parallel in general arrangement, though departures beyond those called for by the raw material itself mark in each case the originality of the individual author. Each description is divided into two parts, a description of the phonemes together with a few texts, and an account of the historical development of the phonemes from the conventional orthographic fiction called Early Modern Irish (roughly 16th century, discounting certain purely orthographic features). The first part in each case comprises, apart from some phonetic preliminaries, chapters on the vowels, diphthongs, consonants, allophonic consonant glides, the phonemic or non-phonemic status of nasalization, sandhi, elision, vowel-length, syllable-division, stress, intonation, and some short texts transcribed phonemically and then into standard Irish spelling, but not translated. Physically the books are uniformly well printed and attractively got up.

The only feature in the general arrangement that is open to criticism is the lack of a section treating the occurring clusters; such statements should be automatically included in any description. In a language like Irish, where morpho-phonemic phenomena are so elaborate and widespread, a chapter on these phenomena would likewise be of great use and interest, but it is not strictly required by the terms of a phonology.

2. Methodology. These books are explicitly based on the 'system of phonetic description' exemplified by Daniel Jones's *Outline of English phonetics* (1936). It is stated that the 'aim is to describe the sounds of the dialect as accurately as possible, either in relation to certain fixed sounds ... or by describing organically their formation' (OC §2).¹ This, one takes it, is the procedure for identifying the PHONETIC or ALLOPHONIC entities. PHONEMES, on the other hand, are defined by a quotation from Jones (Jones §191, OC §29) as 'a family of sounds consisting of an important sound of the language (i.e. the most frequently used member of that family) together with other related sounds which take its place in particular sound-sequences'. An immediate difficulty with this definition is recognized by dB (§10): 'But in a language like Irish where, with regard to vowel sounds, one member of a phoneme is normally as frequently used as another, there is no valid reason for speaking of "principal members" or "subsidiary members"'. The only kind of theoretical statement that will do is of course one that recognizes that there are as many allophones, or members, of a phoneme as there are environments. The fact that we save ourselves time at the same time that we set forth the subphonemic structure of a language by describing gross allophones, as these three works indeed do very efficiently, must not hide this simple truth from us. We get a hint of the methodological trouble this leads to when dB says (§189) that in a certain position 'a front vowel is very slightly more open than it is when between palatals—but for practical purposes it is not regarded as constituting a separate member of a phoneme.' In the definition cited there is further an annoying impreciseness in the phrase 'take its place'. Perhaps a better working statement is to be found in OC §30, where he remarks concerning two different vocoids: 'their occurrence depends on the preceding consonant and neither can replace the other under the same conditions'.

Apart from these specific considerations, there is a further matter that runs through all aspects of these three books. Though it would appear from the above that a division has been perceived and made between the phonetic and the phonemic aspects of sounds, this separation has not in fact been consistently carried out. In the accounts of the separate phonemes no really phonetic transcriptions for the allophones are given at all by way of supplementing the careful verbal descriptions. It would seem sometimes as though 'phonemics' was being made to substitute for 'old-style phonetics' on all levels. Of course, one way of interpreting Jones's definition of the phoneme would make phonemics a sort of additive phonetics.

One example of the confusions to be reaped from a failure to keep these basic matters separate is to be seen in Breatnach's chapter on syllable division. After having remarked on the ease with which the number of syllables may be identified and yet the difficulty one encounters in trying to specify syllable divisions, he says (§334): 'It has been found that for most practical purposes (e.g. in order to mark degrees of stress) it is satisfactory to separate syllables according to some reasonable convention. This method is successfully practised by phoneticians

¹ OC stands for Ó Cúiv and his book; dB for de Bhaldraithe and his book; B for Breatnach and his book.

who are concerned primarily with the teaching of pronunciation. It has been adopted here because no other is possible.'

B continues (§335): 'The result will be found to resemble the scheme outlined by Sommerfelt for Torr. Sommerfelt, however, gives the impression that his syllable division is absolute, implying that, from purely auditory observation, the limits of the syllable may be absolutely defined. No such claim is made for the treatment of the syllable in this work.' I have had the opportunity of discussing this point with Sommerfelt just before finishing this review. Though he firmly maintains his point that observable and describable PHONETIC features occur at points that may be labelled the syllable-border, he remarks that he is well aware that his description of Torr Irish was done a good while before phonemic theory was developed and elaborated. On the phonemic level, Sommerfelt now observes, there is no such thing as syllable division; as in other phoneme-sequence situations, there are simply certain sequences of phonemes, and some combinations occur, while others do not.

B then goes on to discuss the phonetic criteria for peaks of prominence and the difficulty in finding the borderline between such peaks. Because of this difficulty, he proceeds to set up (§343 ff.) some 'conventions' for indicating syllable-division (marked with a hyphen). One of his 'conventions' is that *n'*, *n*, or *l'* at the end of a word, and preceded immediately by another consonant, is syllabic; all such forms are further stated to have alternative forms with *ə* (e.g. *mad'n' ~ mad'en* 'morning'). Since *ə* has already been described (§79) as being extremely short and with various qualities difficult to observe and enumerate, it would seem more in accord with the prevailing patterns of the language to interpret both such forms as containing *ə* (e.g. *mad'en' ~ and mad'en*). The syllabicity itself would then be part of the allophones of *ə*, which already belongs to a structural class of phonemes (vowels) one of whose principal properties may be stated to be, in given sequences including the present type, that of acting as syllable-center. Again (§349), it is stated that in a few cases two consecutive vowels belong to separate syllables; from the examples given it appears that all such cases involve a short vowel following a long /V:V/. If that is so, syllabicity may be stated as subsumed under the allophones of the sequence /:V/.

It would seem then that the phonetic facts of 'syllabicity' might best be stated as functions of the allophonic domain of the stress phonemes. This domain, in the case of Ring Irish, may be stated as the vowel that the stress phoneme clusters with (i.e. is written over), plus any following vowel or length-phoneme (but therefore not a vowel separated from the first vowel by /:/), plus one following consonant provided that a further consonant follows, plus all preceding consonants not already taken up in this fashion by a preceding stress. In West Muskerri, a following consonant goes with whichever vowel bears the stronger stress, or, better put, an unclustered consonant lies in the domain of the stronger adjacent stress. In Cois Fhairrge, an unclustered consonant following length /:/ or schwa lies in the domain of the following stress, otherwise in the domain of the preceding; this regardless of the ranking of the adjacent stresses. This statement cannot be understood to include the situation of /VV/, since adequate data on that sequence are not presented by dB.

In West Muskerry and Cois Fhairrge nasal resonance in vowels, which earlier must have been an important part of the structure of Irish, seems now to be vestigially phonemic largely on an idiolect basis; it may well now structure as a style or personality marker. In Ring nasalization seems to be only allophonic with nasal consonants, yet there is a brief chapter devoted separately to this subject. It is a mark of the youth of our science that people must write treatments (and a good one in the case in question) of features that DON'T occur with such-and-such a status. Perhaps this is because we are still of course writing only partial descriptions of the TOTAL linguistic situation.

If the methodological criticisms advanced above seem not to be put forward very gently, it is certainly not because these books are to be condemned roundly on broad issues; quite the reverse, they are sturdy descriptions and can stand close scrutiny. Honest and valuable descriptions such as these demand honest inspection and criticism. There is, moreover, one overriding consideration that should not go unmentioned. For all the talk about the need for exact and detailed knowledge of dialects, and for all the highly commendable collecting of isolated forms for atlases that has been going on, there is still a critical dearth of full-length STRUCTURED descriptions of individual dialects. The Irish are to be congratulated for being further advanced in this respect than most of us.

3. Restatement of West Muskerry structure. Various fruitful lines of thought are suggested if, on the basis of the detailed and carefully organized data presented by OC, we attempt a restatement of the phonology of West Muskerry. Regarding OC's Chapter II, Principles of Transcription, where phonetic and distributional (i.e. phonemic) matters are intermixed, no detailed comment is necessary here, since the general question of mixing of levels has already been alluded to.

3.1. Vowels. West Muskerry has six vowels (cover-symbol V): /i e a a o u/. In general, these occur freely clustered: with preceding juncture (cover-symbol #) and consonants (cover-symbol C), together with component appropriate to the latter (see below); with succeeding #, C, or length-phoneme /:/; with 'superseding' stress-phonemes (see discussions above and below). Occurring sequences of VV are much more restricted, and are dealt with below. All combinations of the permitted sequence-types occur. Various aspects of the stated cluster-distributions mark off this set of phonemes as a unique structural class, which we call VOWELS.

3.1.1. Vowel allophones. Gross allophones are described in detail by OC for the following environments (cover-symbols are C' for consonant with palatalization, and C' for consonant with velarization or depalatalization):

- | | |
|------------------------------------|-----------------------------------|
| /i/ 1. #iC, C'iC, C'iC' (§59) | /a/ 1. #aC', CaC' (§66) |
| 2. C'i# (§60) | 2. C'a#, C'aC', raC' (§67) |
| 3. iC (§61) (see below for stress) | 3. C'a: (§68) |
| 4. #i:, C'i:C (§57) | 4. see discussion below of schwa. |
| 5. C'i: (§58) | /a/ 1. #aC', C'a#, C'aC' (§69) |
| /e/ 1. #eC', C'e#, C'eC' (§64) | 2. C'aC' (§70) |
| 2. eC', C'e (§65) | 3. #a:C', C'a: #, C'a:C' (§71) |
| 3. #e:C', C'e: (§62) | 4. #a:C', C'a:C', C'a:C' (§72) |
| 4. #e:C', C'e: (§63) | 5. see discussion below of schwa. |

At first glance, and with an eye to reducing morphophonemic alternations if possible, it is tempting to look for complementary distribution in /a/ and /a/, as above stated, in pairs such as /x'a:n/ *Sheaain* : /x'a:n/ *Sheaan*. That this is not possible is, however, demonstrated by pairs such as /m'a:n/ *meadhon* : /s'a:n/ *Seaan*.

- /o/ 1. #oC', C'o#, C'oC' (§73)
 2. C'o#, C'oC', C'oC' (§74)
 3. #o:C', C'o:#, C'o:C' (§75)
 4. #o:C', C'o:C', C'o: (§76)
- /u/ 1. uC', C'u (§77)
 2. #u:C', C'u:#, C'u:C' (§78)
 3. #u:C', C'u:C', C'u: (§79)

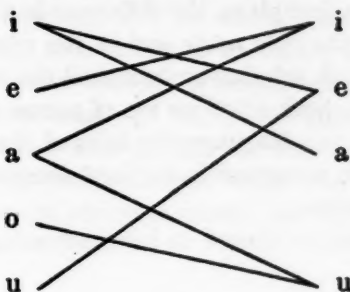
In contrast to the above, OC sets up six separate long-vowel phonemes. Because of behavior noted elsewhere, length is better treated as a separate phoneme; vowels next to length are therefore noted as allophones above.

3.1.2. Schwa. OC set up a thirteenth vowel, which is stated to vary considerably in its allophones; though the precise allophones are not listed and described, the general area is said to be central and rather more open than half-close. The phones subsumed under this caption seem from the examples to be limited to unstressed position. Such sequences of #VC, CV#, and CVC under minimal stress may therefore be put in complementary distribution with allophones of /a/, of /a/, or of both, as noted above.

According to OC, schwa also occurs in stressed diphthongs (see below), but according to his scheme /e/ never occurs in diphthongs. If we consult the accounts of the relevant diphthongs (§86, §92, §94), we find that all the schwa-diphthongs are stated to involve glides in the *e*-area. In VV sequences, therefore, schwa may be considered a bundle of allophones of the /e/ phoneme.

3.1.3. Sequences of VV. Seven 'diphthong phonemes' are set up by OC, and allophonic distributions are described similarly to those for the simple vowels. All the diphthongs are stated to be falling, but the description (§82) includes a bit of unhappy phonetic animism: 'the natural sonority of the more open vowel again makes the starting-point the more prominent element.'

These spans are more efficiently stated as clusters, or permitted VV sequences with stress on the first V. OC's list is *iə, ia, uə, əi, əi, əu, ou*, which I should restate, in the terms elaborated above, as /ie ia ue ei ai au ou/. Amongst younger speakers /au/ and /ou/ tend to fall together (§102); perhaps the resultant sequence is to be interpreted as /au/. If we bear in mind this collapse and rearrange the sequences according to another scheme



it would appear that current change is in the direction of rarifying the already sparse network in the lower part of the unidimensional projection. As soon as the younger speakers no longer have to satisfy the *ou*-speakers, perhaps we may look for a shift back to /au/, thus eliminating /o/ and /a/ from the diagram and tightening the total picture.

Regardless of whether the last point will be right or wrong, if we were to analyze these sequences as unit-diphthongs we should be quite incapable of even contemplating such matters.

3.1.4. Allophonic structure of vowels. Overriding certain general points which could be elaborated from the material summarized in §3.1.1, marked glides occur to and from the base point of a vowel adjacent to a consonant (with component) that has a different tongue position, and most especially when the positions are articulatorily opposite. These glides, which are combined allophones of the consonant, component, and vowel, are described in Chapter VI (§§177-84). They form an important part of the redundancy of Irish.

There are allophonic differences in length which are combined allophones of V, /:/, and #, C, or stress, and which OC outlines in Chapter X (§§221-4).

3.2. Consonants. OC sets up the following forty-two consonants. (For economy, I somewhat compress his articulatory tables.):

b	b'	d	d'	(d')	g	g'			
p	p'	t	t'	(t')	k	k'	tʃ		
v	v'			z	ʒ	ɣ	j		
f	f'			s	ʃ	x	x'	h	
m	m'	n	n'	(n')	ŋ	ŋ'		l	l'
		hn	hn'	(hn')				hl	hl'
								hr	hr'

Some of these elements have vastly more restricted distributions than others. It will be found that by using techniques of rephonemicization, as exemplified, for instance, in Harris's *Methods in structural linguistics* (Chapter 9), the number of elements posited can be considerably reduced, while at the same time certain classes of distributions are broadened and the overall level of distributions is maximized. On the other hand, this operation can be carried out in such a way that very few more unit-symbols are needed to represent a form in this system than are required in OC's system. In fact, from the orthographic point of view, by choosing symbols for the new units that have nonperverse shapes and that yield relatively familiar sequences of shapes, most forms turn out to have almost exactly the same appearance that they do in OC's transcription. In short, to state what is a common analytical experience in matters affecting two solutions worth entertaining in the first place, the difference is not so much in the overall configuration as it is in the cuts made and in the relationships stated between these slices. So long as each solution embodies all the raw data, and hence both are mutually convertible, both solutions are of course equally true. And when a given solution is applied to a language, the facts of the language themselves are naturally never altered; it is impossible to 'do violence to a language' as is sometimes claimed in accusations.

The present restatement is alleged to have advantages over the formulation

of OC, in the terms just discussed, only in the sense that it yields hierarchized statements somewhat more manageably ordered for purposes of writing a complete grammar, and that it discloses within the language structural shapes of somewhat sharper outlines and more compact dispositions. The present restatement also rests, I believed, on a simpler set of theoretical assumptions than does OC's formulation, but to prove that would take us further afield than is proper here. Though all the considerations just mentioned have to do only with the interior properties of the analytical statement, I further believe (but it is founded on little more than faith) that these relationships and their properties are not without connexion with the 'outside world', whatever that is. Finally, it is claimed that the resultant units may be arranged in tabular form more usefully, more suggestively, more 'elicitously' for the total pattern than is true of OC's articulatory tabulation. Such an arrangement will be presented and described below.

3.2.1. Restatement of consonantal segments. The operation here is merely one of resegmenting and reassigning OC's data in ways which are not quite as obvious as was the case with the vowels, though basically exactly the same.

The roster can at once be enormously reduced by extracting a component of palatalization (see below), symbolized *//*, to which we assign a variety of allophonic values. This leaves a still wider variety of residues, which in turn we proceed to assign in this section. The broadest distributional definition that applies to a consonant is this: a consonant (C) is a unit that never clusters with a stress and that may precede a palatalization-phoneme (or component).

We now arrive at segments that may be grouped into units which it will not be profitable to segment further, as follows: /b p d t g k v f z s γ x/ and /m n ŋ l r/, plus /h/. The allophones of these units include the following characteristic features (respectively). For the first group: bilabial plosion, dental plosion, palatal or velar plosion, bilabial friction, tongue-front groove friction, palatal or velar friction; throughout this set there are distinctions of voiceless or aspirated vs. (partially) voiced or unaspirated. For the second group: bilabial nasal resonance, dental nasal resonance, palatal or velar nasal resonance, lateral resonance, retroflex-alveolar resonance; throughout this group as tentatively extracted, voicing is typical of at least part of the segment. For /h/: voiceless glottal and cavity friction. The allophones of these segments are described in general and in particular in the following paragraphs of OC: §§104-9, 112-7, 122-5, 129-33, 135-48; 149 and 49 (for /h/); 152-4, 156, 159-60, 162, 164, 168-74.

The bracketed 'dentals' listed in OC's table (above) refer to alveolar varieties which are used regularly by younger speakers for /d' t' n' l'/ (§§118-21, 126-8, 155, 163); such speakers, however, still use dental varieties in the presence of /r/. These variants are thus structurally to be classed as age-group allophones (or metaphones). OC also notes that his principal informant, an old man who uses pure dentals, has an alveolar /t'/ in /kilt'/ 'quilt', and that the group /lt'/ does not occur in 'native words' (§125). We may then say that the informant has a new cluster in his dialect, a fact which has important implications for the future, but that his old phoneme merely has one more allophone in this environment.

OC's affricate consonant *tʃ* (§150) occurs in no position where it unequivocally parallels only single phonemes; it seems not to occur initially. If we consider it included in the corpus of segments that have just been resegmented, by the extraction of the palatalization component we are left with the residue *tʃ*s. This in turn may be resegmented as a cluster /t/ plus /s/, a sequence comparable to those in /s'ans/ 'chance' and /g'i:ks/, which occurs in *gan glocs ná míocs* 'without a sound'. What is more, the frequency of this span agrees more with that of clusters than with that of single phonemes.

Once the palatalization has been extracted, there now remain unassigned, out of OC's original roster, only *hn*, *hl*, and *hr*. These items appear to be very rare and restricted in their occurrence; the thought of treating them as clusters immediately suggests itself. A canvass of the data, however, shows that some further material is necessary to make a decision possible. One must not be misled by the sequence implied in OC's orthography, with the *h* first; these are simply voiceless resonances with the articulatory position indicated. We therefore seek criteria for establishing which side of the resonance phonemes the clustered /h/ will be assigned to.

Further inspection of the corpus shows that there also occur segments that have already been assigned as partially devoiced allophones of each member of the second group of consonants listed above, namely /m n ŋ l r/, and that these segments, devoiced toward the end of their span, occur before a segment clearly identifiable as /h/. These sequences, then (where R stands for resonance phoneme), are Rh. Some examples, in OC's transcription: *se:r'hu: saothrughadh*, *ar'hi: aithrighe*, *molhəd molfad*, *kail'həd caillfead*, *ku:mhə cumtha*, *l'e:m'həd léimfead*, *du:nhəd dúnfad*, *ki:n'hig' caoinfidh*, *t'aunhəxə *teangthacha*, *taŋ'hən taitneann*. It is clear then that if the voiceless resonances are to be treated as clusters, the sequence can be only hR. Thus, again in OC's transcription: *mə hro:n mo shrón* (doubtless with unjustified orthographic space), *hlog' shloig*, *də hna:v do shnámh*. This sequence is further confirmed by the order of allophones that is found in a more complex type of cluster: *banərhlə ban-altra*, *to:rhn'əx toirneach*.

There is an interesting reciprocal relation between /h/ and /s/ in their clustering properties. They parallel one another in their relatively broad clusterability; but whereas sC is fairly common and Cs fairly rare, Rh seems to be much more common than hR. Note, too, that *hl* % occurs (e.g. *mohl molt*), while **sl* % does not.

OC devotes Chapter VIII to Sandhi. Once palatalization is interpreted as a component, certain features here described become merely allophonic; it appears that palatalization operates over many of OC's orthographic spaces, whatever they are to be interpreted as representing. The balance of the material in this chapter is morphophonemic, and thus should be excluded from the strict phonology.

3.2.2. Summary of consonants. The consonants just extracted may now be summarized and arranged on the basis of their distributional properties.

/b d g v z ɣ/ and /p t k f s x/ oppose one another in point of voicing, and this opposition is reflected to some extent in their distribution. A more fruitful correlation and therefore a more penetrating structural relationship, however, seems

to be established by the distribution of nasals plus stop by point of articulation; the occurrence of /sk/, /sg/, and /zg/, however marginal to the pattern, serves to show the lesser status of voicing correlation.

The structural set R, /m ŋ n l r/, is defined by occurrence with /h/. A subset R₂, defined by occurrence after /h t k g x/, comprises the last three /n l r/.

The set B comprises the labials /b p v f m/, whereby in sB' /s/ always has the velar (non-palatalized) allophone. Likewise set D comprises 'dentals' /d t z s n l/, whereby in rD' /r/ always has the velar allophone. Finally /d t/, isolated by other criteria, form a small class whereby in /xd' xt'/ /x/ has the velar allophone.

A small sibilant class /z s/ is defined by properties of clustering with the stops. The properties of a unique sort shared by /s/ and /h/ have already been pointed out.

/r/ is unique in its occurrence in rhR.

All the above structural properties may be compactly tabulated as follows:

b	p	v	f	m
g	k	ɣ	x	ŋ
<hr/>				<hr/>
d	t	z	s	n
			h	l
				<hr/>
				r

An interesting reappearance of the above patterns on another level is the morphophonemic fact that /b p g k/ alternate respectively with /v f ɣ x/ in the so-called initial mutations, while /d t/ do not so alternate with /z s/.

3.3. Vowel length. The main statements regarding the reinterpretation of this phoneme have already been occasioned by the foregoing discussion. The principal allophonic facts have been recorded in §3.1.1 and in §3.1.4.

So far as I see, the best reason for positing /:/ rather than sequences of like vowels is the fact that in close juncture initial and final single vowels are systematically elided in forms brought into contact with other vowel-terminals (§§206-9); that is to say, VV sequences seem not to occur outside of the limited types already discussed. In West Muskerri I find no evidence for the sequence V:V noted above for Cois Fhairrge, where, by its effect on the allophones of stress domain, /:/ behaves in a limited way like a consonant.

3.4. Palatalization component. This unit occurs in the sequences C'V and C'#. It displays the unique property of having allophonic effect unidirectionally in nonadjacent phonemes; that is to say, except for those cluster-types noted in §3.2.2 (sB', rD', xd', xt'), it affects every C in succession preceding it. There are various phonemic and morphophonemic advantages not only to the positing of this component but also to assigning it to the end of any C group.

Its allophones are various, involving tongue height and point of articulation, depending on the clustered consonants. A special behavior must be noted in connexion with /h/: In the case of that phoneme the tongue height depends according to a statable system (see OC §49 and §149) on adjacent phonemes. We therefore assign occurrences of /h/ on the basis of features which are shared, as it happens, with the environment. A rule on the morphophonemic level is then required to specify that free affixation of // is suspended in the case of /h/.

To clarify the above operation with /h/, let it be said that it is not essential for the positing of // that /h/ be brought into active participation, though it is simpler to solve the question as above than to posit more elaborate clustering rules and allophonic domains, a procedure that would remove /h/ from the category of C, in terms of §3.2.1. Likewise, we must not question the entire phonemic status of a phoneme on the basis of its participation or nonparticipation in the palatalization oppositions, any more than we would question English *t* or *d* because they do not occur before *l* (for many speakers), or Spanish *s* or *θ* because they do not oppose voiced phonemes. This last is the ultimate implication of the remark of J. P. Vinay in his review of these three books, *Word* 3.233 (1947), where he seems to eliminate the affricate phoneme *tʃ* on the argument that it is not opposed by a nonpalatal and 'ne s'entend que dans quelques mots d'emprunt' (despite the fact that he has just cited the Irish sources for the combination on the preceding page). The pertinent argument is of course the cluster consideration already advanced. (A similar cluster criterion must be employed with respect to hR as we have seen; I do not grasp the pertinence of Vinay's arguments on the page cited.)

There is one methodological difficulty which can only be alluded to here, but must be mentioned lest the problem seem to have slipped from sight. It has been the underlying assumption in what has been said thus far that all consonants have an allophonic velar articulation (which incidentally is PHONETICALLY just as marked as the palatal) when not followed by //, and that the addition of // changes the articulation to the palatal allophone. This solution is put forth tentatively as being the simplest, if analytically valid. But it is possible to regard every consonant as automatically clustered with a component (as vowels are with stress), and in that case we also have a component of velarity, even though we might choose to represent it by writing nothing. This possibility raises some interesting problems of assignment in the systems described for the other dialects. If the latter solution were adopted, we should have a suprasegmental sequence running parallel to the segmental phonemes.

Regardless of the implications of the problem just raised, one point is sure: either we have C and //, where C may or may not be clustered with //, or we have C, //, and /', where C is always clustered with one or the other. It would be most uneconomical and, I maintain, phonemically untenable (because it would imply phonemically unassigned segments) to have C, //, and /', where C might also occur alone. The last instance would parallel the setup recently posited by A. Martinet for Old Irish, *Word* 9.1-11 (1953), where he seems to imply the co-existence of /t^l/, /t^v/, /t^a/, and /t/, with the remark (op.cit. 5) regarding initial consonants that 'had there been any consonant coloring there, it would not have

been phonemically distinctive, and actually none is recorded.' Whether it was recorded, especially in Old Irish orthography, is of course one matter; whether in certain positions all possible contrasts occurred is another matter: that is the old question of 'neutralization'. But I firmly maintain that once the opposition had been set up by loss of final syllables (which antedated syllabic syncope), the Old Irish consonants in all positions had to have one quality or another—unless we posit just two components plus an automatic allophone, as above. To maintain otherwise seems to me to imply that members of a system may remain indifferent to that system. Changes in a system revalue the entire system.

3.5. Stress. The allophonic domain of stress has already been discussed above.

Though they are not stated quite so explicitly, there are clearly at least three stress phonemes, which I shall write /^ˈ ˘ ˘/. The evidence is given by OC in §§231–49. Minimal patterns readily establish the primary stress; Vinay is in error in stating (op.cit. 236) that the accent is predictable. The contrast between subsidiary and minimal stress is assured by a pair such as /s'axārd̪:n/ *seachrán* beside /āmādd̪:n/ *amaddn*.

From the evidence given under the rubric of Sentence Stress (§§250–8), it would appear that a fourth stress phoneme is required. This one might be called secondary. The stress structure looks very much like that of English.

3.6. Juncture. Various scattered phenomena seem to point to junctural features, including reference to 'breath-groups' (e.g. in §206). But the material is too meager to warrant any attempt at a systematic statement.

A juncture would seem to be the most economical solution to breaks in the domain of palatalization that occasionally occur (see §47); it is noteworthy that in such cases voicing is also broken. It may also be that juncture occurs in many compounds (§§ 243–9), but this requires more data.

3.7. Pitches and contours. Chapter XIII deals with Intonation; but on the basis of the admittedly summary material presented it is impossible to isolate the pitch phonemes. There appear to be at least three significant levels of pitch. This is the area where competent natives will be able to solve the problem in vastly less time than foreigners, who have no sample frames encoded in their heads to draw on for test purposes, and who thus are swamped at first by the multitude of allophones that depend on the number of syllables and stress sequences with which they interact.

There is a feature called 'special emphasis' (§231, §268, §273) which is hard to pigeonhole; it seems to involve strong stress and marked pitch glides. Is it perhaps a combined allophone of stress /^ˈ/ together with a special contrastive pitch sequence? Or is it perhaps something outside the linguistic system altogether?

3.8. Summary of nonsegmental phonemes. Besides vowels (V) and consonants (C), we have set up vowel length /:/, four stresses, component(s) /'/, at least one juncture /#/ , and several pitches. Whether stresses and junctures structure together as they do on English is not known. Otherwise these seem to be separate structural classes of phonemes as listed.

4. Remarks on Cois Fhairrge. The following comments point up the differences between the dialect of Cois Fhairrge and that of West Muskerry, just described.

The vowels, restated, are /i e æ a o u/; each of these occurs with and without length /:/ . According to dB (§§58-61), the vowel /a:/ occurs only long. We must take dB's word for it, though Éamonn Mhac an Fhailigh, in his review of this book in *Éigse* 5.143 (1946), remarks that others have noted some instances of the same quality as short. On the other hand, schwa is reported (§§83-90) as a lone short. I regard these two qualities as allophones of a single phoneme.

The following vowel sequences are given by dB: /iə ei uə ai au/ and /uəi uəu/; in these, schwa may of course be replaced by /a/. Stress is on the first member of every sequence.

If we apply to dB's table of consonants the same techniques as those used on the West Muskerri data, we reduce his forty-eight items to the following scheme; note the different status of // in this dialect:

b	p	w	f	m	} with or without //			
g	k	ɣ	x	ŋ				
d	t		s	N				
<hr/>								
(d	t)					} only with //		
j	č	n l						

5. Remarks on Ring. This is the latest in date of the three books; though it follows exactly and explicitly the general plan of the others, it has certain minor departures. The chapter on Stress is much richer in morphophonemic information than the corresponding chapters in the other books, but the phonological side of the chapter is weaker.

The vowels are /i e a o u ə/. Since it occurs restressed in this dialect (§80), /ə/ is phonemic. /a/ is not included in B's chapter on vowels, but it must be considered since it occurs in English loans which are stated (§§215-6) to be part and parcel of the everyday vocabulary. Vowel sequences are /iə ia uə ai au əi əu/.

The thirty-four items in B's consonant table may be restated as follows:

b	p	v	f	m	} with or without //			
g	k	ɣ	x	ŋ				
d	t		s	n				
			h	l				
<hr/>								
j	ç			r	} only with //			

Abweichende spät- und vulgärlateinische Perfektbildungen. By FRANK G. BANTA. (Diss. Bern.) Pp. xiii, 124. Freiburg in der Schweiz: Paulusdruckerei, 1952.

Reviewed by ROBERT L. POLITZER, *Harvard University*

The purpose of this work is to collect (from a great variety of miscellaneous sources) and classify Latin perfects which deviate from the traditional Latin standard, defined by the author as the usage of the classical Latin authors of the

1st and 2nd centuries A.D. The author also hopes that his book will give a picture of the breakdown of the Latin perfect forms and the simultaneous development of the Romance perfects.

Like all investigators of Late Latin material, Banta is faced with a difficult methodological problem: the straddling of the synchronic and diachronic points of view which is inherent in the subject matter. On the one hand, Late Latin must be approached as a coherent synchronic system; but on the other hand, the deviation from the standard is usually of greater interest to the investigator. Moreover the author of the Late or Vulgar Latin text was under the synchronic influence of the actual popular speech of his period and under the diachronic influence of the classical Latin standard.

Banta solves this problem by a synchronic-diachronic classification. The perfects are classified synchronically according to the type of perfect formation: A. simple perfect formed by the addition of the perfect ending: *spondi*; B. Schwundstufe perfect: *puli*; C. Umlaut perfect: *fici*; D. reduplication perfect: *spondi*; E. u-perfect: *fregui*; F. s-perfect: *morsi*. A somewhat similar type of classification, utilizing small letters, indicates which category the perfect belonged to according to the classical Latin standard. Finally there is a second synchronic classification according to the stem from which the perfect is formed: 1. present stem: *compell-i*; 2. stem which can be either present or perfect: *prand-id-i*; 3. perfect stem: *compul-s-i*; 4. stem which is neither present nor perfect: *funx-i*. A form like *relinqui* is thus classified as A1c, *spondidi* as D1d, *morsi* as F1d. This scheme of classification makes it clear that the new formations in the Late Latin perfects are due to shifting around between perfect classes rather than the creation of new morphological categories.

After establishing this scheme of classification, Banta arranges his perfect forms accordingly, gives an alphabetical enumeration of his sources with a list of the forms found in each work, and finally alphabetizes the collected forms with references to the works in which they are found.

As a result of this arrangement, Banta's monograph impresses the reader as a reference work rather than a study which attempts to formulate specific conclusions. While Banta occasionally discusses the historical development of forms, there is no systematic attempt to correlate the findings of the work with the Romance outcomes. Banta applies his scheme of synchronic classification quite mechanically to a period of some 1000 years. He does not distinguish between his sources. As in any historical problem—and the topic of the work is essentially of historical interest—the material could have been presented in the form of a succession of synchronic stages. If Banta had subdivided his material into smaller periods, we could perhaps have gained a picture of the dynamics involved in the shifting among perfect classes. It might then have been possible to evaluate the frequency of the new formations statistically, and to establish significant differences of frequency in time or space. Which of the perfect types becomes most productive? Are there any semantic types which tend to become associated in the same perfect class? These are among the questions which one would like to see answered and which Banta never raises.

Yet Banta's study is undoubtedly very useful. Perhaps in a doctoral disserta-

tion the systematic collection of facts is more important than the discussion of underlying problems. At any rate it is to be hoped that Banta himself will try to examine the theoretical, historical, or semantic implications of the material which he has collected with so much diligence and accuracy.

A unidade da România ocidental. By THEODORO HENRIQUE MAURER JR. (Universidade de São Paulo, Faculdade de Filosofia, Ciência e Letras: Boletim 126, Filologia Românica no. 2.) Pp. 227. São Paulo, 1951.

Reviewed by ROBERT A. HALL JR., *Cornell University*

In the development of historical linguistics, as elsewhere, first things have had to come first. This has meant an emphasis, at the outset, on phonology and inflection, in order that we might have a sound basis for tracing structural changes in the blood-and-bone, as it were, of language. But this necessary emphasis has often resulted in an imbalance, in that there has been a tendency to neglect other aspects of linguistic history, especially in the customary introductory textbook with its almost exclusive presentation of 'phonology and morphology'. Yet word-formation, syntax, and lexicon are certainly as deserving of our attention as are phonology and inflection: the latter are indeed more basic but also more skeletal, whereas the former give a language its distinctive physiognomy and individual characteristics.

Maurer's book is a discussion of these relatively neglected aspects of Romance structure, with a view to demonstrating their importance in the history of the western part of the Romance-speaking world.¹ His main thesis is that Western Romance (= Italo-, Gallo-, and Ibero-Romance) constitutes a special unity in the Romance family, not so much because of common features of phonology and inflection going back to 'Vulgar Latin', as because of later developments in word-formation, lexicon, and syntax which represent common borrowings from medieval Latin and reciprocal borrowings among the Romance languages themselves from the Middle Ages onwards (particularly, however, from French into the other languages).

In support of this thesis, Maurer presents his material in three main sections: Os fatores da unidade românica ocidental (16-56), O vocabolário românico ocidental (57-146), A gramática românica ocidental (147-215), followed by a short concluding section (216-21) and a bibliography of works consulted (222-4). The first section is concerned primarily with cultural history, and outlines the relevant intellectual conditions of medieval times in three chapters: Unidade de pensamento e cultura da România ocidental (18-26), with subsections treating of the Roman Church and of certain unifying social institutions (merchants,

¹ As has become customary in continental European languages, Maurer uses the term *a România* to refer to the territory where Romance languages are spoken. This is a very convenient term, and one wishes it could be borrowed into English, in some such form as 'the Romania'. Unfortunately, this would cause difficulty, both because of possible confusion with the name of the country *Ro(u)mania* and because it is contrary to English syntactic habits to use the definite article with a proper name of this kind; we therefore have to keep to the more awkward but more idiomatic expression *(the) Romance-speaking territory* or *world*.

minstrels, universities, and monks); *O latim—lingua universal da Idade Média* (27-44); and *O prestígio cultural do Francês na România medieval* (44-56). The section on Western Romance vocabulary covers somewhat more than we usually include under this heading, since it not only has chapters dealing with loanwords (60-76) and loan-translations (76-84), but also one dealing with word-formation (84-146, the longest single chapter in the book) and containing subsections on suffixation, prefixation, and composition. The third section begins with some very brief observations on phonology (149-50) and orthography (150-1), and a relatively short chapter on morphology (152-65), the major part of the section being taken up by a consideration of syntax (165-215).

Maurer's basic procedure is to examine a given item of vocabulary, derivational element, or construction, with regard to its traceability to 'Vulgar Latin'² and its geographical distribution, and to present conclusions as to its primarily 'popular' or 'learnèd' nature. He deals almost exclusively with material from the Western Romance literary languages and from standard Roumanian, and his customary test for learned origin in doubtful cases is whether the element under consideration is found in Roumanian or not; if it is not, that is held to clinch the case for learned origin, since Roumanian was separated from the rest of the Romance-speaking world after the fall of the Roman Empire. There are two objections to this type of argument: (1) that it is an *argumentum ex silentio*, and (2) that it leaves out of consideration both the Western Romance dialects and also Sardinian, which is even more archaic than Roumanian. As a result, Maurer often goes too far in assuming that the absence of a given element from Roumanian indicates that it must be of learned origin in the rest of Romance; for instance, he considers as primarily erudite such suffixes as *-icus* (96), *-ālis* (96-7), *-ānus* (97-8), *-īnus* (98-9), *-ē(n)sis* (104), *-āta* (108-9), and *-aticum* (111-3), and the verbal prefix *ad-* (121-2), all of which show little or no vitality in Roumanian but are preserved in Sardinian,³ a fact that points clearly to their having been present in Proto-Romance.

Furthermore, it is questionable whether we are really justified in placing so much emphasis on vocabulary and syntax as the prime unifying factors in Western Romance, at the expense of the more basic aspects of phonology and inflection. There are also a number of phonological and inflectional isoglosses which set off Western Romance from both Roumanian and Sardinian; we may mention such features as the 'Vulgar Latin' seven-vowel system; the 'West Romance sound-shift';⁴ the fusion of 'VL' /č/ and /c/ into /c/ in Gallo- and Ibero-

² Maurer's use of this term is not clearly defined, but in general (as with many other scholars in the field) seems to refer to popular Latin of the Imperial period, without a clear distinction between this and the ancestral source of all the Romance languages (Proto-Romance by definition).

³ For these suffixes in Sardinian, in the same order as given above, cf. M. L. Wagner, *Historische Wortbildungslehre des Sardischen* 15-7, 37-40, 50-2, 52-5, 85, 92, 91, and 134-5 (*Romanica Helvetica* No. 39; Bern, 1952).

⁴ For convenience, we keep the term 'West Romance sound-shift', even though the shift did not spread to absolutely all of the West Romance territory; it is characteristic of Western Romance, and is certainly absent from Central and Southern Italy, Roumanian, and Old Sardinian as an indigenous development.

Romance; the reduction from three to two or one in the number of thematic vowels in the imperfect tense-sign in Gallo- and Ibero-Romance; and the development of a third verbal stem (the so-called 'future' stem) in addition to the two ('present' and 'preterit') that go back to Proto-Romance. We normally consider phonological and inflectional isoglosses to be more important than those in the fields of derivation and lexicon, because the former occur in large numbers of lexical items; Maurer seems to have neglected this factor of relative importance. The syntactical aspect of Romance studies has, in general, been neglected, primarily because until very recently we simply have not had any adequate means for analyzing syntactical structure, to say nothing of making a comparative reconstruction of Proto-Romance syntax. Maurer discusses certain types of more or less learned syntactical constructions, but here again seems to neglect the basic types of phrase- and clause-structure which form an underlying unity beneath later accretions.⁶

In short, Maurer's thesis is only partially valid. In its positive aspects, it is of course quite justified: Western Romance is indeed unified in its vocabulary and syntax, as compared to Balkan Romance and Sardinian. But if this were all, we would not be justified in setting up Western Romance as a separate unity in the Romance world. (English, despite all its derivational, lexical, and syntactical borrowings, is still not a Romance language!) Phonology and inflection still are more important than syntax and lexicon in determining linguistic relationships, and they go hand in hand with these latter in establishing Western Romance as a major grouping.

Nevertheless, Maurer's book is very valuable. It can serve as a useful counter-balance to the prevailing overemphasis on nothing but phonology and inflection, and will be of use to students of all the Romance languages, since its content is of cross-cultural validity. Maurer has made a careful selection of material from a number of sources, which will be very helpful as a repertory of the Western Romance languages' borrowings from Latin and from each other.

Studien zur Entwicklungsgeschichte des Frankoprovenzalischen. By HELMUT STIMM. (Akademie der Wissenschaften und der Literatur [Mainz]: Abhandlungen der geistes- und sozialwissenschaftlichen Klasse, Jahrgang 1952, Nr. 6.) Pp. 160. Wiesbaden: Franz Steiner Verlag, 1953.

Reviewed by ROBERT A. HALL JR., *Cornell University*

In this monograph, one of the latest efforts to define and classify that elusive entity 'Franco-Provençal', Stimm treats certain phonological characteristics of the region, from the point of view of their chronological sequence. After a

⁶ Maurer's use of the term *syntax* seems to refer principally to the more complicated aspects of syntactic construction. On p. 166 he cites G. Pepe's curious remark 'We do not know, above all, whether this sermo plebeius had a syntax of its own [just how could a language exist without a syntax of its own?] but perhaps it had only the simple syntax of the ignorant people that spoke it.' The only way to make any sense out of this remark is to assume that by the term *syntax* Pepe means something special, over and above the basic patterns of phrasal and clausal combination common to both the sermo plebeius and more elegant usage.

brief introduction (5-11), three main chapters treat, respectively, the development of stressed mid vowels and *au* (12-63), and of the stressed vowels followed by *u* or *ɥ* (64-93) and before *i* or *j* (94-142). A summary of Stimm's conclusions (142-8) is followed by a full bibliography (149-53) and an index of Latin etyma (154-60). Stimm's treatment is thorough, detailed, and careful. As material, he uses not only present-day attestations (in the ALF and monographic treatments of individual dialects), but also two 13th-century sources, one of which (the *Somme du Code*) is earlier and phonologically more unified and conservative than the other (a collection of legends and miracles).

Stimm reaches the conclusion that a number of the features usually considered as linking Franco-Provençal with northern Gallo-Romance (e.g. diphthongization of stressed mid vowels in free syllables) are of relatively late origin, and that, at an earlier stage, Franco-Provençal was much closer to southern Gallo-Romance—a conclusion with which there can be little quarrel. Stimm refuses (145-7) to accept von Wartburg's hypothesis of a Burgundian superstratum as the cause of the earliest differentiation of Franco-Provençal; instead, he suggests the autonomous diffusion of the development $\acute{a} > \acute{e}$ in free syllables after palatal and other traits from Lyons as a cultural center. He suggests a historical connection between Franco-Provençal on the east and Poitevin—which shows some of the same traits—on the west; according to this theory, both would be surviving relic areas of a very early (4th-5th century A.D.) focal area centered on Lyons. Unfortunately, this theory does not seem to have any firmer historical basis than von Wartburg's; and the similarities between Franco-Provençal and Poitevin can perfectly well be explained by assuming that both are transitional areas and reflect the uneven spread of North French features southward on either side of the Massif Central. In this connection, Stimm (like various others) has badly misunderstood your reviewer's position in *Lg.* 25.1-14 (1949); but the place to discuss this is in a separate article.

Vocabulario dei dialetti della Svizzera Italiana. Ed. by SILVIO SGANZINI. (A cura della Repubblica e Cantone del Ticino con l'aiuto della Confederazione Svizzera e con un contributo del Cantone dei Grigioni. Fascicle 1, *a-agnesa*, pp. xxxvi, 40, with 2 maps in flap. Lugano: Tipografia La Commerciale, 1952.

Reviewed by ROBERT A. HALL JR., *Cornell University*

This is the long-awaited first fruit of the undertaking begun by Carlo Salvioni in 1907, designed to provide the Italian-speaking part of Switzerland with a dialect dictionary comparable to the *Schweizerisches Idiotikon*, the *Glossaire des patois de la Suisse Romande*, and the *Dicziunari rumantsch-grischun*. The Swiss-Italian dictionary has had a chequered history, passing from Salvioni's direction to that of his pupil Clemente Merlo in 1920, and in 1936 to that of the latter's pupil Silvio Sganzi. Since 1940, the offices of the *Vocabulario* have been established in Swiss territory and the undertaking has received official support and subvention; editorial work has resumed progress, to the point where publication has now begun.

In the first fascicle of the *Vocabulario* are contained the 'front matter', the

first forty pages of text, and (separate from the book itself) two outline maps of Italian Switzerland. A brief 'Presentazione' by Karl Jaberg (v-vii) is followed by a more extensive 'Prefazione' (viii-xi) in which Sganzini sets forth the history of the work. The 'Introduzione' (xiii-xxxvi) explains various technical features of the *Vocabolario*: its aim and structure, and the organization of the individual articles (xiii-xv), the transcription (xv-xvii), the regions and communes represented (xvii-xxii), an extensive bibliography (xxii-xxxv), and a list of abbreviations (xxxv-xxxvi).

It is evident from this fascicle that the *Vocabolario* will, if carried to completion on the present scale, rank as one of the foremost dialect dictionaries of the Romance world. The articles are elaborately organized, with the key words in conventional spelling but accompanied by phonetic transcription, and with indication of regional variants, definitions, citations, and etymology. There are occasional block cuts (five in this fascicle) of local artifacts or relevant illustrations from ancient manuscripts, church paintings, and the like. The work has evidently been prepared with great care, especially in regard to the distinction of the various meanings and grammatical function of words, and in its remarkable wealth of citations. The etymological discussion includes not only local attestations and distribution, but also the relation of words to the rest of the Romance territory. For instance, the discussion of medieval Poschiavese *accolla* 'group of isolated houses' (23) affords valuable further evidence for the earlier relation of the Italo-Alpine and Rhaeto-Romance dialects with northern Gaul.

This undertaking has far more than narrowly regional significance, for both its content and its method, and therefore deserves all the support we can give it, especially by individual and library subscription.

Essai sur la phonétique du parler rhétoroman de la Vallée de Tavetsch (Canton des Grisons—Suisse). By LÉONARD CADUFF. Pp. xiv, 229 (photo-lithoprinted). Bern: A. Francke AG Verlag, 1952.

Reviewed by ROBERT A. HALL JR., *Cornell University*

The Tavetsch valley is the westernmost part of the Grisons, containing the upper source of the Rhine; it is bounded on the north and west by Canton Uri and on the south by the Val Leventina in Canton Ticino. Because of its marginal position among the Rhaeto-Romance dialects, we might expect the Tavetsch to be of especial interest; yet Caduff's book, a Lausanne dissertation, is the first extensive treatment it has received. The Tavetsch reveals a close affinity to the neighboring valley on the east, the Médels; by and large, the speech of these two valleys is closer to that of the Engadine than it is to that of points down the Rhine in the Surselva, such as Disentis (186-8). In some respects (e.g. diphthongization) the Tavetsch preserves an earlier stage than Disentis, where, according to Gartner (whom Caduff follows), the Rhaeto-Romance diphthongs were simplified by a regression to the stage of single vowels. In other respects, the Tavetsch and Médels valleys seem to have been influenced by the Ticinese dialect of Val Blenio (with which the Médels region is connected by the Lukmanier pass): e.g. *-driu* > *-fyr*; *-ice* > *-aš*.

Caduff's book is a thorough treatment of the historical phonology, of the traditional type, taking the phonemes of 'Vulgar Latin' as a point of departure, and discussing the representation of each in the sounds of modern speech. A brief introduction (1-19) is followed by treatments of vowels, first stressed (23-58) and then unstressed (58-68), and consonants (71-177); after a brief recapitulation and conclusion (179-89) come extensive indices (189-224), a table of contents (225-6), and an outline map of the Tavetsch and neighboring valleys. The dialectal examples are given in narrow phonetic transcription, with no indication of the structural ordering of the various sounds; modern dialectology should analyze the distribution, not only of raw phonetic material, but also of sounds in their functional classifications. The photo-lithoprint reproduction of the typescript is not easy to read, especially since the dialect material has all been inserted by hand; there are a fair number of misprints, and occasionally diacritical marks have been left off.

This monograph represents a good beginning in the investigation and analysis of the Tavetsch dialect; presumably it will be followed by more material, in historical morphology and vocabulary. Further desiderata would be a good historical treatment of syntax (for which descriptive techniques now afford us the means to establish combination-types synchronically and to reconstruct them historically) and—if funds should be available—an extensive collection of texts. Even the fullest treatment of linguistic structure is like a textbook of anatomy more or less richly illustrated with plates; but a collection of texts is like a motion picture of a living person.

La pronunciación en el español del Valle de México. By JOSEPH MATLUCK. Pp. xxvi, 125. México, D. F., 1951.

Reviewed by ISMAEL SILVA-FUENZALIDA, *Universidad de Chile*

Matluck's work is intended as an additional study in the determination of the dialect areas of Spanish in the New World. It is clear, however, that both in theory and in method the work fails to satisfy modern requirements for a study of dialect geography. In this respect it would be useless to discuss its many inadequacies in detail. For example, the author uses Navarro Tomás' *Cuestionario lingüístico hispanoamericano*, which is notorious for its outdated methodology, and fails to cite such a standard work as Kurath's *Handbook of the linguistic geography of New England*. Even so, Matluck does not achieve the results obtained by Navarro Tomás in *El español de Puerto Rico*. The author presents no cultural analysis to validate his age groups and social stratification groups, nor does he provide the reader with any indication as to the degree of social mobility in an area so close to Mexico City. The particular history of each of the localities studied (Xochimilco, Texcoco, and Tlalnepantla) presumably has not been investigated. Matluck gives as his reason for selecting these communities (vii): 'Elegimos estos lugares por ser cabeceras de municipios que tuvieron, y en parte siguen teniendo, una fuerte población indígena y que, por lo tanto, conservan rasgos típicos de la región.' He fails, however, to restrict his choice of informants

to speakers unaffected by prolonged residence in outside areas or other communities.

Since Matluck's analysis is not presented in standard form, with mapping of the data, but with only vague and sporadic references to particular geographic areas and social strata, this book constitutes, at its best, only a very general description of the pronunciation of Spanish in the Valley of Mexico as a whole. This fact, added to the author's unawareness of structural linguistics in general and of phonemics in particular (cf. xvii), renders his study of very limited value.

Os estudos de linguística românica na Europa e na América desde 1939 a 1948:

Suplemento bibliográfico da *Revista portuguesa de filologia*, Vol. 1. Ed. by MANUEL DE PAIVA BOLÉO. Pp. xi, 521, [24]. Coimbra: Caso do Castelo, 1951

Reviewed by LAWRENCE B. KIDDLE, *University of Michigan*

In the late forties, when European scholars began to recover from the disastrous effects of the war on their activities, one sign of this recovery was the preparation of synthesis volumes in various fields. These volumes, which appeared in the first years of the present decade, sought to re-establish the lines of international scholarly communication broken by the war, as well as to present summaries of progress made during the period of isolation. In the general field of linguistics the UNESCO-CIPL publications, *Bibliographie linguistique des années 1939-1947* (2 vols.) and *Bibliographie linguistique de l'année 1948*,¹ are already well known. Sever Pop's *La dialectologie*,² widely used and esteemed by linguists, can be considered in part characteristic of the 'synthesis' trend of postwar European scholarship. In the narrower field of Romance linguistics we have Alwin Kuhn's *Romanische Philologie: 1. Teil, Die romanischen Sprachen*.³ The present note concerns another volume of synthesis, published in Portugal as a bibliographical supplement to the *Revista portuguesa de filologia*. Since this volume seeks to accomplish the same purpose as those just mentioned, we shall describe it and evaluate it with reference to similar volumes.

The first important characteristic of this work, one that makes it unique among the comparable volumes just referred to, is that it is the cooperative work of seventeen contributors. The work was planned, directed, and organized by Manuel de Paiva Boléo. The motivating idea behind the present work, as carefully stated by the editor in his preface, was to provide something much more useful to investigators than a bibliographical repertory like the UNESCO-CIPL *Bibliographie*. Such a repertory provides a list of titles that prove so tempting to researchers that they often spend time and money locating and purchasing books of little real research value. The seventeen contributors were asked to review critically the field of Romance linguistics from a geographical standpoint: each one was charged with reviewing the activity in his own area during the period covered. All works of importance in each area were to be mentioned, with

¹ Reviewed by Hall, *Lg.* 26.185-6 (1950), 27.610-1 (1951).

² Reviewed by Hall, *Lg.* 28.119-22 (1952).

³ Reviewed by Malkiel, *Lg.* 28.509-25 (1952).

analyses of and commentaries on the outstanding productions. Within these norms the contributors had complete liberty of selection and arrangement.

Obviously a volume with such an organization will show great variation in the separate contributions, and will lack the tight logical organization of a one-man analysis like Kuhn's or Pop's. The only advantage offered by a cooperative volume is its more intimate and more complete coverage. In the present volume the following areas are treated by the following collaborators: Sweden (Lars Wiberg), Denmark (Poul Høybye), England (W. D. Elcock), Switzerland (Heinrich Schmid), France (G. Gougenheim), Argentina (Ana M. Barrenechea and Narciso Bruzzi Costas), Belgium (L. Warnant), Canary Islands (Juan Régulo Pérez), Spain (A. Zamora Vicente), Catalonia (R. Aramon i Serra), United States (E. B. Williams), Holland (E. C. van Bellen), Germany (Alwin Kuhn), Brazil (Serafim Silva Neto), Spanish America (M. L. Wagner), and Italy (Ruggero M. Ruggieri). Several things are noteworthy about this list of areas. The absence of Portugal is explained in the preface by the editor, who states that Portuguese activity will be chronicled in the second volume of the work. There is no explanation, however, for the omission of the Balkan Romance area.⁴ It is very striking that the complete work will include seven Ibero-Romance areas (Spain, Portugal, Catalonia, Argentina, Spanish America, Brazil, and the Canary Islands). One of these areas, the Canary Islands, appears for the first time in a selective bibliographical commentary of interest to linguists. Contributions vary in length from the regrettably brief article on the United States (9 pages) to the unduly extended one on Italy (76 pages); the average contribution contains approximately 30 pages. Another feature worthy of special mention is the inclusion of necrologies. There are twenty of these, varying from a short paragraph to several pages; in ten cases there are full-page portraits of the deceased scholars. There is an index of authors (479-97), a list of abbreviations (498-501), a topical index (501-8), and a word index with listings by the separate languages (508-15). The investigator who uses this work has ample avenues of approach to its contents.

Fifteen of the seventeen articles follow the space-time limitation imposed by the editor.⁵ The two that depart from it are those by Régulo Pérez on the Canary Islands and by Wagner on American Spanish. Both of these articles treat whole fields and refer to all important modern works.⁶ A few of the contributors merely list selected titles and offer either no descriptive comment or, at best, only a few sentences summarizing the content of the work (Elcock, Williams, Kuhn); others fortunately give a fairly complete summary or analysis

⁴ Alwin Kuhn, op.cit., devotes forty-one pages (116-56) to Balkan Romance, with references to such linguists as Pușcariu, Caracostea, Capidan, Papahaghi, Constantinescu, Daicoviviu, Procopovici, Petrovici, Iordan, Cazacu, and Iroaie. To these we may add Bogdan, Boršč, Budagov, Densusianu, Galdi, Gazdaru, Křepinský, Olteanu, Rosetti, and Dragănu. See UNESCO-CIPL *Bibliographie 1939-1947* 1.119-20, 2.385; *Bibliographie 1948* 116-9.

⁵ Apparently there was some confusion about the period to be covered. The title of the book names 1939-1948; but many of the contributors stop at 1947 (Schmid, Barrenechea and Bruzzi Costas, Warnant, Zamora Vicente, Aramon i Serra), and one goes to 1949 (Wagner).

⁶ This leads to duplication, especially between Wagner's article on American Spanish and the Argentine summary by Barrenechea and Bruzzi Costas.

of each work selected (Wiberg, Høybye, Zamora Vicente, and, above all, Schmid and Wagner). Some of these lengthy analyses approach full-scale reviews in value, e.g. Wagner's comments on Henríquez Ureña's *El español en Santo Domingo* and on Tomás Navarro's *El español en Puerto Rico* (375-80). Several contributors list only scattered critical reviews of the individual titles or none at all (Høybye, Elcock, Gougenheim, Williams, van Bellen, Warnant, Ruggieri, Barrenechea and Bruzzi Costas); others record a large number of reviews for most of their titles (Wiberg, Kuhn, Schmid, Régulo Pérez, Wagner, Aramon i Serra, Zamora Vicente) and thus provide the investigator with another way to ascertain the research value of a title that attracts him. Most of the contributors give only an objective description, but a few offer critical comments of the type suggested by the editor (Warnant, Schmid, Régulo Pérez, Zamora Vicente). Many have a valuable introductory statement summarizing linguistic studies in their areas before 1939. From Høybye, Warnant, Schmid, and Ruggieri the reader receives an explanation for the preponderance of descriptive studies in Sweden, of dialectal and comparative studies in Belgium and Switzerland, and of the influence of Ascoli, Gilliéron, and Croce on contemporary Italian linguistic scholarship. For Spain the years prior to 1939 were turbulent; both Zamora Vicente and Aramon i Serra describe the success of attempts to carry on the work of such prerevolutionary institutions as the famed Centro de Estudios Históricos, and to renew the publication of such reviews as the *Revista de filología española*. The majority of the contributors use a Romance language (French 7, Spanish 4, Portuguese 3, Italian 1) for their articles. The only exceptions are the British and American contributors, Elcock and Williams, who use their native English.⁷ The separate divisions of the articles vary, of course, with the trends of research in the country in question. Such standard classifications as Phonetics, Phonology, Morphology, Syntax, Lexicography, Etymology, Dialectology, Linguistic geography, Onomastics and toponomastics, Versification, and Textual criticism appear frequently. The most satisfying classification, from a reader's standpoint, is that employed by Alwin Kuhn in his article on Germany, with essentially the same divisions as his larger bibliographical summary and commentary, *Romanische Philologie*. A glance at the rubrics listed above suffices to show that much more is included than what we know as linguistics; the title of the volume is somewhat misleading. Several of the individual contributors must have sensed this, since they use the term Romance philology instead of Romance linguistics in their titles (Wiberg, Aramon i Serra, Williams, van Bellen, and Neto). The most completely satisfying article of the entire book is the Swiss summary by Heinrich Schmid, mainly because he follows more closely than any other contributor the plan suggested by the editor in providing selected titles, well-rounded analyses of the individual works, ample critical reviews for each title, and his own critical comments on certain selections. To all of these he adds an author index covering his own contribution.

A striking impression comes to the reader when he compares the productivity

⁷ Both Elcock and Williams could and should have used a Romance language. Most of the contributors of non-Romance extraction (Wiberg, Høybye, van Bellen, Kuhn, Wagner) use a Romance language instead of their native tongues.

of German and Italian Romance linguists with that of British and American scholars as reflected in this volume. Even allowing for possible exaggeration on the part of the German and Italian contributors, the meagre listings of British and American productions is enough to dismay and annoy the reader of this work. Elcock explains convincingly why so little work by British scholars appears for the period and why this work is mainly in the field of French. The article on American research, however, gives an erroneous impression of inactivity. Of the unduly brief presentation of American material, only nine pages in all, approximately six pages are devoted to critical editions of texts; the rest is given over to scattered titles chosen capriciously and lacking in descriptive or critical comments. It is not sufficient to refer the reader to the annual bibliographies of such American journals as *PMLA*, *Modern language quarterly*, and *Italica*: one of the contributors to this very same volume, E. C. van Bellen, complains bitterly (300) about how difficult it is for European scholars to obtain copies of foreign periodical publications because of currency exchange complications; he is of course referring principally to the United States. Manuel de Paiva Boléo planned this publication to give scholars like van Bellen an easily accessible instrument for satisfying the bibliographical needs of their research. In view of this aim, Williams should certainly have been more inclusive, more complete in his selections and more detailed in their presentation.

Referring in detail to the last pages of Williams' article (282-4), we find a brief paragraph of eight lines covering the field of etymology. Reference is made to the importance of studies by R. A. Hall Jr., Yakov Malkiel, and Leo Spitzer; one article by Malkiel is cited. The last paragraph of the article, labeled 'Miscellaneous', mentions the 'distinguished work' in certain defined fields by Pierre Delattre, Francis Millet Rogers, T. Navarro Tomás, R. A. Hall Jr., Victor R. B. Oelschläger, Norman P. Sacks, Dorothy C. Clarke, and R. S. Boggs. Many more details should have been given concerning this work, especially because of its 'distinguished' nature. But these are not all the researchers who did meritorious work in the period 1939-48. What about J. H. D. Allen Jr., A. Alonso, E. C. Armstrong, L. D. Bailiff, D. L. Bolinger, W. E. Bull, J. Corominas, J. E. Gillet, L. H. Gray, E. F. Haden, Anna G. Hatcher, H. R. and Renée Kahane, H. Keniston (one article of his is mentioned, 284), R. Levy, C. H. Livingston, H. F. Muller, M. A. Pei, J. B. Rael, C. C. Rice, and R. K. Spaulding? Many of these did work far more significant than some that is listed in Williams' inadequate catalog. In fact, the decade 1939-48 saw several publications of major importance by American Romance linguists. The following all deserve descriptive summaries of the type that Schmid does so well: T. Navarro Tomás, *Manual de entonación española* (New York, 1944); id., *Estudios de fonología española* (Syracuse, 1946); M. A. Pei, *The Italian language* (New York, 1941); Madaline W. Nichols, *A bibliographical guide to materials on American Spanish* (Cambridge, 1941); Victor R. B. Oelschläger, *A medieval Spanish word-list* (Madison, 1941) and R. A. Hall Jr., *Bibliography of Italian linguistics* (Baltimore, 1941). Although Williams may have had some of these publications in mind when he wrote his laudatory references to specific individuals, they deserve much more than that. The few items that Williams does list seem capricious. Thus, he lists

three articles by C. E. Kany, together with a brief mention of his important *American-Spanish syntax* (Chicago, 1945). It is hard to understand what led the compiler to choose just these three articles out of the seven that Kany published in the *Hispanic review* in 1943-5. And why are there no references to dissertations, especially to those that have been published? This is one of the valuable aspects of the contributions by Wiberg and Schmid.

The book is attractively printed with clear type, satisfactory margins, and good paper. There are unimportant printer's errors, especially in those articles that appear in English; but in general, titles in all languages and bibliographical details are carefully recorded. This care contributes to making the work one of outstanding value of Romance linguists throughout the world.

Os estudos de filologia portuguesa de 1930 a 1949: Subsídios bibliográficos. By GIACINTO MANUPPELLA. (Instituto para a Alta Cultura: Publicações do Centro de Estudos Filológicos, No. 4). Pp. 246. Lisboa: Centro de Estudos Filológicos, 1950.

Reviewed by LAWRENCE B. KIDDLE, *University of Michigan*

This volume modestly claims to provide 'some data for a future critical bibliography of Portuguese philology'. Actually it provides a great quantity of carefully tabulated bibliographical information in a field where such assistance will be highly appreciated. The volume seeks to list all titles dealing with Portuguese philology that appeared in the period 1930-49, together with significant reviews of the works listed. The date chosen for beginning the book's time span was that of the publication of Antenor Nascentes' important *Dicionário etimológico da língua portuguesa*. The work was originally planned to appear serially in an Italian journal, but its length made such publication difficult. This was fortunate, since it is clearly preferable to have the work in a single volume. The change in publication plans explains why the notes accompanying most of the bibliographical entries are in Italian, whereas the title, the indices, and the section of additions and corrections are in Portuguese—a fact which will of course cause no serious difficulty to the readers for whom the book is intended. The combination of Italian and Portuguese is handled accurately in the printing of the book, and the work is characterized by unusual care in the printing of names and titles in other languages as well.

The publication contains 1894 numbered bibliographical entries. Since an entry can appear more than once under separate numbers (for example: these pairs of items refer to the same work: 400/706, 537/538, 141/517), the total number of individual items is less than the figure given. There is no clearly defined system of cross references to care for the necessity of repeating items under the different divisions of the book. Entries are classified under the following headings: Journals; Bibliography; Bio-bibliography; General works; Homage volumes and collections; Historico-linguistic questions; General and special dictionaries; Historical and descriptive grammar; Orthography; Phonetics and phonology; Morphology; Syntax; Style, literary language, and language esthetics (i.e. stylistics); Dialects; Etymology and word history; Lexicology, Lexicography, and semantics; Technical, scientific, and professional language; Slang

and cant; Onomastics and place names; Brazilian aspects of Portuguese; Linguistic relations with other peoples and orientalia; Literary and documentary texts and various sources for the history of the Portuguese lexicon. Additions and corrections are appended to the main body of the text (199-208). A glance at the divisions of the book is enough to reveal the excessive and arbitrary compartmentalization carried out by the compiler. The difficulty that this procedure causes in the case of a title that belongs in two or three of the narrowly split fields listed above is considerably lessened by the index of materials and subjects treated and by the index of words studied. There is also an index of authors.

The typical bibliographical entry gives the following data: author, title, full publication details, page dimensions (of books), and, for articles, the complete name of the journal, with the volume, year, number of fascicle, and pages. Somewhat unusual in extensive bibliographical repertories like this is the insistence on spelling out the complete name of each periodical every time it is cited, instead of using the commonly accepted abbreviations. The author makes note of this in his introductory words and apparently he considers it an improvement in the preparation of bibliographies. Readers will agree with the present reviewer in admitting that such a practice does noticeably facilitate the use of the bibliography. Almost all entries are followed by short descriptive notes about the contents of the work or, in the case of extended polemics, about related publications. While these notes are never critical in any sense, they are extremely valuable to the reader. There follows in chronological order a list of significant reviews of the title in question. There is no indication whether the review is favorable or not, as there is in Hall's *Bibliography of Italian linguistics*.

In order to ascertain the value of this bibliography compared to similar compilations in the same field, your reviewer checked about two hundred random titles in this and three other recently published bibliographies—CIPL *Bibliographie linguistique des années 1939-1947* (2 vols.), CIPL *Bibliographie linguistique de l'année 1948*, and Woodbridge and Olson's *A tentative bibliography of Hispanic linguistics* (Urbana, 1952). The volume under review is clearly the easiest of these to consult. It offers more bibliographical information and lists more reviews of individual titles than the other works. It offers more titles than Woodbridge and Olson, and, while the abundance of titles it presents is comparable to that of the CIPL volumes, it exceeds the latter in accuracy. After its ample coverage of the field the greatest excellence of the present book is precisely this consistent accuracy. In numerous cases where errors appear in one or more of the other volumes referred to, there is a gratifying regularity of careful annotation in Manuppella's work.

Studier og tydninger. By JOHS. BRØNDUM-NIELSEN. Pp. xi, 240. København: J. H. Schultz Forlag, 1951.

Reviewed by W. P. LEHMANN, *University of Texas*

This volume of reprinted articles by Brøndum-Nielsen, published between 1910 and 1949 in various journals and *Festschriften*, was issued by his friends in celebration of his 70th birthday. The collection will be useful primarily to

specialists in North Germanic dialects: most of the articles deal with highly restricted areas in the development of Danish, chiefly in lexicography but also in phonology and syntax; there are essays also in stylistics and literary history, and one on the influence of typesetters on the language of the 16th century. Brøndum-Nielsen's arguments for the new etymologies he proposes, or for his interpretation of various linguistic developments, are impressively supported; witness the fifteen-page discussion of the relationship between Modern Danish *Haandtage* and *Haandtave* 'slap on the hand', or the eight and a half pages of examples in the treatment of inverted conditional clauses following the apodosis. His conclusions seem to have been generally accepted.

Possibly the greatest interest of this Festschrift is the difference it illustrates between the aims of a young linguist in 1910 and those of a linguist today. In 1910 the basic problems were solved; linguistic methodology had been established; the chief outlines of the history of the Germanic languages had been clarified by Grimm, Brugmann, Paul, and others; it remained to fill in the details, and to write more extensive grammars based on fuller collections of the data, as Brøndum-Nielsen did for Old Danish. Today we have no such assurances. Nineteenth-century methodology has been overturned by the work of Saussure, Trubetzkoy, Bloomfield, and others; but no one new method has been agreed on or seems capable of universal application. And we lack the most fundamental studies for historical linguistic work, as may be illustrated by the numerous articles appearing on the sound systems of Gothic, Old English, and other languages, and the smaller number dealing with the morphological or syntactic systems. While descriptions that seem adequate to us in methodology and information are being worked out, one of the more fortunate products would be the continued appearance of scholars with Brøndum-Nielsen's erudition and his ability to fit together minor facts.

Dativstudien: Dativus sympatheticus und dativus comparisonis in der norrönen Sprache. By VEMUND SKARD. (Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo; 2. Hist.-filos. Klasse, 1951, No. 2.) Pp. [viii], 139. Oslo: Jacob Dybwad, 1951.

Reviewed by STEFÁN EINARSSON, *Johns Hopkins University*

These are two interesting studies on the use of the dative in the Scandinavian languages, chiefly Old Norse, i.e. Old Icelandic and Old Norwegian, but with side glances at Old Danish and Old Swedish as well as at the modern dialects and standard languages.

The first study, on the dativus sympatheticus (DS), is rooted in Wilhelm Havers' *Untersuchungen zur Kasus-syntax der indogermanischen Sprachen* (1911); it tries to verify or correct his statements about Old Icelandic, which were based on sketchy material and are hence liable to revision. To begin with, Skard disagrees with Havers in defining the phenomenon, limiting his DS to the first three (or two) of Havers' six classes. According to Skard, the meaning of the verb refers (1) to the human (or animal) body or its parts; (2) to the human soul; or (3) to everything belonging to a person. Skard gives the frequency of the three

groups as approximately in a ratio of 100 : 10 : 1. This is why he suggests dropping group 3 altogether.

Skard notes that the DS can alternate with a prepositional phrase or a genitive. Thus, theoretically, we can have: *spjót stóð í brjósti Ólafi* or *spjót stóð í brjósti-nu á Ólafi* or *spjót stóð í brjósti-(nu) Ólafs*. Since all these phrases are translated 'A spear stood in Olaf's breast', I have in my *Icelandic* (110-1) defined the form *Ólafi* as a possessive dative. Skard naturally did not use my book, but if he had known it, he would have found Blöndal's statement about Modern Icelandic confirmed: the DS now often belonged to a rather literary style, while the prepositional phrase is the most colloquial usage.

Havers found that instances of the DS were chiefly pronouns of the first and second person, but Skard shows that this does not hold for Old Norse.

Skard devotes a long chapter to *þiðreks saga*, in which he finds a characteristic variation of the three constructions. He believes that this variation marks the German-Norwegian linguistic milieu in Bergen, and that it can best be explained by assuming that the author wrote stories from oral narrative rather than translating from German texts.

He finishes this part of the work with a list of phrases which he has not found in dictionaries. Under *blóð* his construction *e-m stöðvar blóð* is misleading; it should be *e-r stöðvar e-m blóð*.

In discussing the dativus comparationis (DC) the author's starting points are two recent studies: E. Benveniste's *Noms d'agent et noms d'action en indo-européen* (Paris, 1948), and H. Mörland's *Ablativus und quam* (Oslo, 1948). Benveniste combines the dative construction with the (comparative) suffix **-yes-*, and the particle construction with the (comparative) suffix **-tero-*. He shows that the comparison by case is a closer unit, often presupposing identity of the compared parts, while comparison by the particle (*en* 'than') is a looser construction. Scholars had long noticed that the case comparison was more common in poetry than in prose, but Mörland explains this as being due to the style of poetry, which favors concentration.

Skard divides his material on case comparison into seven categories; he treats first poetry, then prose. For particle comparison he sets up eight groups. The other Scandinavian languages (Old Danish, Old Swedish, Mod. Icelandic, and Mod. Norwegian) are more briefly treated. Of his examples illustrating case comparison I shall quote one as belonging to a type which has gone unnoticed by scholars: *v menn ... æða fim* (dat.) *fleiri* 'five men ... or more than five'.

Under his case constructions I looked in vain for light on a peculiar point of similarity between Middle English and Modern Icelandic, a point which long ago caught my attention. In the *Ancren riwle* (in Emerson's *ME reader* 200) we read: *Nenne weopmon ne chasti ze, ne ne etwiteð him of his undeau, bute zif he beo þe overcuðre*. In Icelandic: *engum karlmanni skuluð þér hegna, né ávíta hann fyrir ósið sinn, nema hann sé því nærgöngullí*. In English: *chastise no man, nor blame him for his bad manners, unless he is familiar (impertinent) to a high degree*. In other words, a dative (or instrumental) of comparison with a comparative here means, both in Middle English and in Modern Icelandic, 'a very high degree'. I may add that I think the construction in Icelandic is possible only after

nema 'unless'. There is one example which Skard quotes from *Völsunga saga* (68.8-10) that may be comparable: *ræð ekki slíkt við mig, nema þú sért hverjum manni fremri* 'do not discuss this with me, unless you are better than any other man (= a very outstanding man)'.

Vokalismen i Iddemålet. By REIDAR MYHRE. (Skrifter frå Norsk Målførearkiv, No. 1.) Pp. iv, 118. Oslo: Jacob Dybwad, 1952.

Reviewed by EINAR HAUGEN, *University of Wisconsin*

There is a tendency in many countries to regard dialectology as a branch of study distinct from linguistics, and many dialectologists show a fine disregard for the results of linguistic science. An example of this is the vast collection of Sever Pop at Louvain, at least if we trust the review in *Lg.* 28.119-22 by Robert A. Hall Jr. Another example is the present monograph on the vowels in the dialect of Idd, Norway, which is fairly typical of what is being done in Norwegian dialectology. One reason may be that dialect monographs are often written by beginners in linguistic study, who use their own local dialect for cutting their linguistic eyeteeth. It is disappointing that so few of the professors in charge of dialect study have taken the trouble to revise their methods in keeping with the general advance of linguistic science. The present monograph offers nothing of methodological interest, and would therefore not be worth reviewing in *LANGUAGE* except for two things: (1) it is the first in a new series of monographs published by the Norwegian Dialect Archives in Oslo (Sigurd Kolsrud in charge), and (2) it offers a chance to state some of the weaknesses which are inherent in the traditional approach to language of such monographs.

This particular dialect is interestingly located in being the southernmost Norwegian dialect along the Swedish border, surrounded on three sides by Swedish speech. The author commendably devotes a good deal of attention to comparing its forms with those of the neighboring communities, and finds more Swedish elements than in the other Norwegian dialects of the area. The features of difference are unfortunately merely enumerated, as *disjecta membra*, and no attempt is made to evaluate them with respect to frequency or structural position. It is impossible to gain a general perspective of the nature of these differences: some are merely phonetic, others are phonemic, still others distributional; the author even includes in the same lists some that are purely morphological. There is a good deal of useful information about the geographical and administrative relationships of the community. We are told that today the dialect is in rapid dissolution and that it was almost impossible to find speakers of 'genuine' dialect; unfortunately we learn nothing about the kind of changes resulting from this breakdown of dialect.

The body of the monograph is historical in nature. What does it tell us of the present-day dialect? Only that certain changes in the quantity and quality of vowels took place from Old Norwegian times to the present. There is no synoptic table of these changes, nor any attempt to set up a structural picture of them, only a plodding vowel-by-vowel account of what happened to each one. Like other East Norwegian dialects this one has retained its long-vowel system quite

well, aside from the general increase of labiality in the back vowels. But the Old Norwegian diphthongs have been monophthongized, while new diphthongs have developed from spirant /g/. The short vowels offer a complex picture, with a general but not completed change in the direction of lowering and fronting. At least two entirely new phonemes have developed, an /ä/ derived from short /a/ and /e/, and an /ʌ/ from short /o/ and /ɔ/, sometimes from /a/ and /ø/.

It is hard to be sure of this, however, because of the sketchiness with which the sounds of the dialect are described and the complete lack of any distributional study. This is not to be blamed on the author, who was only following the models available to him; but it is to be hoped that the time may come when higher demands will be made on the descriptive side of dialect study. Since the present study was made without benefit of any phonemic theory whatever, it can only be judged on the basis of its phonetics, which is poor. Three skimpy pages are allotted to a 'phonetic survey', which is essentially a description of the phonetic symbols used to transcribe the dialect.

The vowels are set up in a tabular form devised by Bell and Sweet, as revised by Norwegian phoneticians to suit their own needs. Although the originator of this practice, Johan Storm, as far back as 1908 deplored the uncertainty of meaning attached to the terms 'wide' and 'narrow', they have persisted in Norwegian dialect description to this day, after having been generally abandoned elsewhere. In the table of the present monograph they clearly mean nothing more than the difference in quality between long and short of the same vowels. They are therefore superfluous, since the information they give could be conveyed by a simple statement to the effect that short vowels are laxer or lower than long vowels. Eliminating this double set of columns would simplify and clarify the description of the vowels, which otherwise is in terms of the standard high-mid-low, front-mixed-back, and round-unround of English phonetics.

But within this framework there are a number of odd placings. The sound /a/ is placed in the mid-back-unround space, /o/ in the mid-back-round. One wonders how the author arrived at this conclusion, since Storm, to whom the author constantly refers, describes the former as low-back, the latter as high-back, which agrees much better with the reviewer's observation. In his verbal description of /a/ he says he means the 'usual East Norwegian a' and of /o/ he says it is the 'usual closed Scandinavian o'. This kind of description is worse than useless, since it assumes that there is a prevalent kind of East Norwegian *a* or Scandinavian *o*. When the tabular description contradicts the verbal one, we are confused indeed.

There are certain special sounds in the dialect, however, which have caused him even greater difficulty. One of these is the /ä/, which is derived from /a/ and sometimes from /e/. It is verbally described as the 'broad æ-sound', and it must certainly lie somewhere between /a/ and /æ/. Yet it appears in his table in the high-back-unround space! In a note he says that perhaps it 'could have been placed as front-low'. It is hard to understand how any one could be in doubt about whether a sound is at one end or the other of the linguistic spectrum, but the same doubt is expressed by Ingeborg Hoff (*Skjetvemdlet* 10 [Oslo, 1946]), who has much more experience in dialect research than our author. She finally

decided to place it in the low-front space, and there does not seem to be any reason to question this decision. Another sound (or group of sounds) which causes difficulty is the one written with a special α digraph having no lower hook on the e ; we shall here reproduce it as $/\Delta/$. Myhre is uncertain whether this should not have been written $/\ddot{o}/$, his short, lowered $/\phi/$, but decides to describe it as low-front-round. In almost the same spot, but low-mixed-unround, he has an $/\nu/$ which he finds difficult to place; perhaps it should have been low-mixed-round.

Some of these uncertainties are the result of phonetic inexperience. Others are due to the complete absence of phonemic considerations in the choice and arrangement of the symbols. His table includes altogether (if we disregard the length marks) 17 different vowel symbols. Several of these are quite obviously superfluous, since they are clearly indicated to be allophones of the others. Thus he places a dot over the $/y/$ and the $/u/$ which occur as the second element in diphthongs; inconsistently enough, he does not use a special symbol for the $/i/$ that occurs in the same position. A short $/\phi/$ occurs in the table, but is nowhere used in the text, since its place is everywhere taken by $/\ddot{o}/$. The latter symbol is unnecessary, however, since it differs from the long $/\phi/$ in exactly the same way as other short vowels from the corresponding longs, and will be covered by the above-mentioned statement about this difference. Then there is a low-mixed-round vowel written $/\alpha e/$ which occurs only before 'cacuminal' $/l/$. The information given is not sufficient to determine whether this should be regarded as an allophone of $/\phi/$ or of a phoneme which will be discussed in a moment.

This is the phoneme which was reproduced above as $/\Delta/$. That it is a genuine phoneme is evident from the fact that it occurs in all positions, and that it is possible to find minimal pairs in the author's material, although he does not bother to reassure the reader by presenting them to him. But the other low sound which he writes $/\nu/$ has a much more doubtful position. He does not seem certain, as noted, concerning its phonetic value. He claims that it is different from $/\Delta/$ and $/\ddot{o}/$; he finds $/\Delta/$ to be 'deeper and more rounded'; just how one low sound can be 'deeper' than another is also puzzling. His reasoning here is not convincing, and is completely overthrown by his own practice in writing the words containing these sounds. In the phonetic survey he says that he has decided to write $/\nu/$ instead of $/\Delta/$ in words like $/b\ddot{o}lle/$. He does so on pages 15, 23, and 25; but on page 51, where he treats the sound change of $/o/ > / \Delta /$, he uses the symbol $/\Delta/$ not only in *bolle*, but in a whole flock of other words of exactly the same type! The fact that he has forgotten to change these spellings, if that is what happened, shows that it probably doesn't make much difference. A word of a different type, *golw* 'floor', is written with $/\Delta/$ on page 51, as is a whole group of other words of the same type, while the word is written with $/\nu/$ on pages 94 and 268. On page 69 the two phonemes are handled as if they were long and short of the same vowel, both being derived from ON $/a/$. This minute analysis of symbols is undertaken to show what could be gained in accuracy of notation if the author had made sure that his symbols reflected only significant differences in his material.

All in all we have succeeded in reducing his symbols from 17 to 12, and of

these one, /ə/, occurs only in unstressed position and could probably be regarded as an allophone of /e/ or /ɛ/. This makes a system of 11 vowels, or the usual East Norwegian nine plus the two extra phonemes /ä/ and /Λ/, which are also found e.g. in the dialect of Elverum and other neighboring dialects. In nearby Skjetve, on the other hand, according to the report of Ingeborg Hoff, /ä/ and /ɛ/ appear to have coalesced. The vowel system thus groups itself into four levels of opening (high, mid-raised, mid-lowered, low) and three of rounding (spread, round, over-round):

i	y	u
e	ø	o
ɛ	(ə)	ä
ä	Λ	a

If the writer had had some such picture clearly in mind, he could not have allowed himself to confuse his description by the continued use of the terms 'opening' and 'lowering' to describe some of the historical changes of the old short vowels. These hallowed terms in Norwegian dialectology (invented by Amund B. Larsen in 1885) have recently been attacked by Ernst W. Selmer, and quite rightly. The words themselves are badly chosen, since they do not correctly describe the phonetic facts, and it may be questioned whether the distinction between them is valid. 'Opening' includes such various changes as /i/ > /e/, /u/ > /ø/, /e/ > /ɛ/, /a/ > /ä/, /o/ > /ə/; but a number of these are also 'lowerings' and hence indistinguishable from the former (cf. Larsen, *Oversigt* 25-6). An example of the confusion in this monograph is the statement that before /pp/ the vowel /a/ is 'lowered' (40); but how can a low vowel be lowered? The sound law /u/ > /u/ (54) is described as a 'lowering' also; a moment later the same word is used to describe the change /u/ > /o/. It appears to be used simply as the opposite of 'opening'. Larsen described his 'opening' in terms that made it sound like a centering movement; but actually it is a combination of lowering and fronting, and had best be described as consisting of these components.

In conclusion, it should be emphasized that a dialect is also a language and deserves to be treated accordingly.

Studies in honor of Albert Morey Sturtevant. [Ed. by L. R. LIND.] Pp. [v], 169. Lawrence: University of Kansas Press, 1952.

Reviewed by ARCHIBALD A. HILL, *Institute of Languages and Linguistics*

Albert Morey Sturtevant is one of the giants from an earlier and perhaps happier day when no one would have thought of a division between language studies and literature. His contributions to Scandinavian and Germanic linguistics and to Scandinavian literary history form a monument to both disciplines, testifying to the fact that they were once happily wedded under the name philology—and may be again united when both pursuits have reached greater maturity. The present graceful volume is a tribute to that union, past and future.

Only one of the articles, that by Adolph Benson, 'The problem of Catholic sympathies in Swedish romanticism', is completely without linguistic content,

and this article, by its demonstration of how the Catholic-mystic tendencies of German romanticism were worked out in literary and philosophical expression in a community where Catholicism was unacceptable, gives—if not linguistic perspective to literary matters, what may be as repaying, namely wide sociological and cultural implications. The other articles range from bibliographical (Didrik Arup Seip, 'On the original of the Codex Regius of the Elder Edda' finds evidence of a Norse prototype) through folkloristic (Jess H. Jackson, 'Melkólfs Saga ok Salomons Konungs', presents a fragmentary manuscript tale with text, translation, and a commentary on the folklore motifs contained in it), to the field of questioned documents in Erik Wahlgren's linguistic and epigraphical consideration of 'The runes of Kensington'. Of those articles that are not strictly linguistic, this is one of the most interesting. It is perhaps not altogether free from the argumentative quality which so frequently creeps into writings on the most famous and most questioned of American Norse antiquities—now that the Newport Tower has faded into obscurity. The line of argument is that most of the peculiarities of the inscription might be explained by drawing on varied and complicated sources of error, change, and survival, but that if the assumption is made that the stone is a 19th century forgery, all peculiarities are immediately explainable in terms of a single source. The argument is an impressive use of the criterion of simplicity.

Easily the most valuable of the linguistic articles is 'The impact of English on American-Norwegian letter writing', in which Einar Haugen pursues his researches on bilingualism and borrowing. The article is drawn from a closed corpus, a group of 289 letters to an American foreign-language newspaper; it thus shares with the recent book by Charles C. Fries the distinction of analyzing a body of language material where the data are definite and verifiable. Haugen again presents his analysis of the process of borrowing, and once more explodes the notion that languages in contact can reach a stage where it is impossible to say which one is being used.

Kemp Malone's article 'The phonemes of Icelandic' is remarkable as one of the very few instances in which a scholar who earlier presented an elaborate phonetic account of a language has taken the trouble to bring it into line with current principles of description. The article presents a very clear picture of distributional facts, but employs an unusual terminology (a cluster of two differing consonants is called a 'twosome') and is marred by certain inadequacies of phonemic analysis. Thus (9-10), 'The tectals [i.e. velar-palatal nasals] do not occur in initial or final position, but the vibrant [i.e. voiced] may occur in sandhi if the next word begins with a tectal [i.e. velar-palatal, whether nasal or stop]. Example: *Jón Grímsson*. In such cases, however, the vibrant is hardly a member of the tectal phoneme but serves as an allophone of [n].' This statement involves the notion of an entity which can be a member now of one phoneme now of another, and so violates the truism that phonemes do not overlap. The alternation of entities is here rather a matter of morphophonemics than of phonemics proper. Another crux of phonemic analysis is the following problem, which the paper presents, but does not solve. Long and short vowels occur; so do long and short consonants. But the long vowels occur only before short consonants, while

the short vowels occur only before long consonants and clusters. The presence of consonant clusters which act in the same way as long consonants makes the setting up of long consonants as separate phonemes uneconomical. Such longs can be written as clusters of same consonants, and the long and short vowel distribution can then be regarded as purely complementary. Such a solution would have the result of reducing the four types of vowel nucleus set up by Malone to two, an advantage of some importance since languages with four vowel nucleus types seem to be rare at best.

Stefán Einarsson, in 'Compounds of the *mann-skratti* type', gives an interesting collection of this construction from Scandinavian languages and even from Sanskrit. The compound, he says, is capable of being rendered into English only by the periphrasis *a devil of a man*. The type is rare in English, but not unknown. The construction *man mountain*, though perhaps not written with a hyphen, is surely an example of the same formation.

Caroline Brady, finally, studies 'The synonyms for *sea* in Beowulf'. Her conclusion is that the words and phrases used in connection with the sea are quite sharply differentiated in meaning, and contain literal rather than conventional or figurative reference. The approach throughout is rather like the definitions given in a dictionary, but these definitions seem backed by careful consideration of the contexts in which the terms occur. The paper is a demonstration of how much there is still to be learned about Old English stylistics, and a proof once again that literary students can contribute to linguistics, just as so many of the papers not only of this volume but of Sturtevant's own bibliography demonstrate that linguistics can contribute to literary study.

Shakespeare's pronunciation. By HELGE KÖKERITZ. Pp. xv, 516. New Haven: Yale University Press, 1953.

Reviewed by ARCHIBALD A. HILL, *Institute of Languages and Linguistics*

A work on a subject which has been investigated before must be an improvement in method and results if it is to succeed. *Shakespeare's pronunciation* passes this test with flying colors. The previous work was Wilhelm Viëtor's;¹ it was characterized by such absurdities as the transcription of most unstressed vowels as spelled, and the belief that many of Shakespeare's sounds were so similar that he could rhyme them together in spite of differences. In more positive ways, one of the great virtues of Kökeritz' book is that it represents a thorough examination of the whole Shakespeare corpus, marshalling all the evidence from spelling, puns, and rhyme. In contrast, Viëtor's study was based primarily on the poems alone. Kökeritz presents the first complete rhyme index, itself a work requiring years of patient labor. Further, there is a kind of question which he answers in an extremely authoritative manner, as a result of the thoroughness of his investigation. The question can be formulated in this way: What modern phoneme should we assign to a given word in order to approximate most closely Shakespeare's pronunciation of that word? It is an important question, which

¹ *Shakespeare's pronunciation*, two volumes: *A Shakespeare phonology* and *A Shakespeare reader* (Marburg, 1906).

perennially interests most students of Shakespeare. They have a practical and urgent interest in approximating Shakespeare's pronunciation, and at the same time an equally practical and urgent interest in using a system of pronunciation which will be understandable to hearers who are not experts. The question has literary bearings which force themselves on anyone who has tried to direct a Shakespeare production, or even to read one of the plays aloud. It is therefore a high compliment to Kökeritz' transcriptions that they recommend a pronunciation which not only does justice to individual words but also presents a system like Shakespeare's, and at the same time understandable to speakers of Modern English.

Since it is the first full-length treatment of this phase of early Modern English in nearly half a century, *Shakespeare's pronunciation* deserves an accurate and detailed examination, not merely aimed at its improvements over its predecessors, but also at the system of Shakespearian pronunciation it presents, and at the assumptions and methods on which its results rest. The book is one in which such matters are not described with detailed explicitness, but have to be inferred from discussion of individual problems. As a result the examination which follows will necessarily be one which employs many specific and perhaps minute examples, since only through these is it possible to arrive at the larger picture of system and method.

I shall begin with the system of sounds. Kökeritz uses such words as phoneme and allophone throughout the book,² but does not choose to organize his sounds in these terms. It is therefore a tribute to the excellence of his research and the clarity of his presentation that his summary of sounds (340-2) allows the reader to reorganize the sounds into a phonemic system with probable justice to Kökeritz' intentions. In this summary he gives a total of twenty symbols for vowels and diphthongs. These can be organized into a phoneme of length and seven basic vowel phonemes, with uncertainty arising only about the eighth. Kökeritz' statements about length are explicit, and leave no doubt that he wishes to set it up as a phoneme.³ The basic vocalic phonemes are then these:

1. A high front phoneme /i/, with an allophone [i] before length.
2. A mid front phoneme /e/. This had no higher allophone before length, but had an opener allophone [e] before length plus /r/.
3. A low front phoneme /æ/. This had no allophonic variation before length, and it is important to note that the occurrence of length with /æ/ was itself conditioned (or partly so?): 'followed by [f, s, θ] in *after, ask, bath*, etc., and before

² Often with considerable looseness, as in this sentence (210): '... the early orthoepists ... were unable to analyze a phoneme like [ju:] and its allophones ...' Throughout this review, the symbols used in phonetic notation are so far as possible those used by Kökeritz. The principal departure is that phonemic symbols have been enclosed in slant lines.

³ Thus 6: '[Shakespeare's English was characterized] by a marked quantitative distinction between historically long and short vowels, e.g., [i:] in *beet* and [ɪ] in *bit*'. And 161: 'It is very important to remember that the quantity or relative length of a vowel was a significant characteristic; though the short vowels were probably somewhat different in quality from the corresponding long ones, particularly [ɪ] and [ʊ], the main distinction between the pairs [æ] and [æ:], [e] and [e:], [ɪ] and [i:], [ʊ] and [u:], [ʊ] and [u:] was that of length. The modern tendency in certain types of English to replace a combined qualitative and quantitative distinction by that of quality alone, which seems to characterize American English, was unknown in eNE.'

n + consonant in Romance words like *dance*, *grant*; also for ME *au* in *laugh*, *half*.

4. A mid central phoneme / Λ /. It had an allophone [ə], which occurred in unstressed position and in the diphthongs [əɪ] and [əʊ]. It had a further allophone, somewhat higher, [ɜ], which occurred with length before weakly articulated [ɹ] in *bird*, *burn*, etc.

5. A high back phoneme /ʊ/, with an allophone [u] before length.

6. A mid back phoneme /o/, which occurred only before length.

7. A low back phoneme /ɒ/, with a higher allophone [ɔ] before length.

8. The uncertain phoneme is that in low central position. Kökeritz sets up several varieties of sound, an [a:], an [a], and an [ɑ:], all in this general area. The [a:] was conditioned and regular in words with older *er*, *ar* where it occurred before weakly articulated [r]. In *after*, *ask*, *bath* it appeared as a substitute for /æ/. The [a] was a substitute for the low back in *want*, *wash*, and the [ɑ:] was the corresponding long before weakly articulated [ɹ] in *war*, *warm*, etc. These sounds were all dialectal variants in origin, and it might therefore be argued that they had no phonemic status. It is Kökeritz' opinion, however, that they existed in London speech side by side with the forms for which they were substitutes. They would therefore have to be set up as an independent phoneme, the norm of which was /a/, but which had the allophone [ɑ] after /w/.

The picture which now emerges is worth giving in diagrammatic form:

I		U
ɛ	Λ	o
æ	a	ɒ

Allowing for differences in the choice of symbols, this system is startlingly like the one arrived at for contemporary English by a number of phonemicists often thought of as radical. The similarity extends even to the limitation on /o/, which is now also limited to long forms, or (as it is now generally written) to the nucleus /ow/, at least in its more common occurrences.

A number of points in this system invite discussion. One of these is the absence of a high central phoneme, the modern /i/. The existence of this phoneme in modern English is often denied, and there may indeed be dialects (or idiolects) in which the sound does not exist, or in which it exists only in complementary distribution with other sounds. Yet for many varieties of speech (including my own) there can be no doubt that /i/ occurs under stress, and in contrast with / Λ /, /ɪ/, and /e/. The usual key word is *just* in *just a minute*, which contrasts with *just* in *a just decision*, and also with *gist* and *jest*. Part of the trouble over this phoneme is that the writing system of English has never provided a way of indicating it unambiguously. It is perhaps for this reason that the phoneme was isolated only very late, though the lateness of recognition cannot be taken as proof that the sound itself is a recent development. There are, of course, indications of the sound in writing, such as the spellings *jist* and *jest* for *just a minute*;⁴ another possible indication of the sound is the form which appears in northern ballads as *mither*, and still another is the modern (often southern American) *whup* for *whip*. In short, alternations in spelling between *i*, *e*,

⁴ This development in Modern English is referred to by Kökeritz on 211, but he interprets the dialect forms only as [dʒɪst] and [dʒɛst].

and *u* can be taken as indications of the possible occurrence of /i/. Under the circumstances, there can hardly be conclusive proof of the existence of this phoneme in Shakespeare, but there are a number of places where Kökeritz' researches give the possibility of /i/. For what they are worth, I will list them, with some discussion.

114: Kökeritz discusses the line from *As You Like It* 1.3.20, 'cry hem and have him'. There is, as he indicates, a pun on the indistinct *h'm*, and I would agree that 'cry hem' might mean something like 'cry the marriage banns'. If so, the spelling suggests the modern distinction in many dialects between /əm/ 'them' and /im/ 'him'.

129-30: *Hamlet* 3.2.262 has a variant reading *mistake* : *must take*. Kökeritz suggests a pun on the basis of [ə]. But /i/ in *mistake*, with possible stress for the sake of the pun, seems at least possible.

190: 2 *Henry IV* 2.1.31 gives Mrs. Quickly's *Lubbars head*, for *leopard's head*. The orthodox explanation of such variations is given by Jespersen,⁵ who traces such forms to a dialectal ME mid front round vowel, which after unrounding, gave the variants [i], [e], and [ʌ]. Preservation as /i/ would equally well explain the variant spellings. Kökeritz also takes up a variant reading from *Romeo and Juliet* 5.3.170 *rest* : *rust*, where *rest* is usually chosen by modern editors, though *rust* is possible. Kökeritz suggests the possibility of a dialectal development of *rest* to /rast/ in the neighborhood of /r/. If so, it is an example of a centering tendency such as has often produced /i/. On the other hand, it is just possible that the spelling *rust* here means /rist/, a development possible from /e/ in any dialect which had not yet developed /ʌ/ from /u/.

210-1: The material here presents a much stronger possibility of /i/. Kökeritz discusses spellings in *i* and *y* from French /ü/, the usual representative of which is *u*.⁶ As instances of such spellings he cites *Gillian* for *Juliana* (*Comedy of Errors* 3.1.31); the striking *misicall* (*Two Noble Kinsmen* 1.3.76), in which to a modern ear /i/ seems particularly likely, and the even more striking pun *Jupiter* : *gibbeter* (*Titus Andronicus* 4.3.79-85). The pun, of course, is not exact, but seems more likely if we assume /i/ in one or both words. The basic reason why this list of forms, not all of which I have quoted, seems to me strongly suggestive of /i/ is not only that the history of the group is one of the most likely sources for the phoneme, but that I have heard /i/ in words of this class in Modern English.

Even more important than the possibility of a high central vowel, however, is the formulation which sets up length as a phoneme. Kökeritz is certainly orthodox in doing so, since though most of them are pre-phonemic, all standard handbooks of early Modern English assume the existence of long and short vowels, usually in such a way as to make it clear that length is phonemically significant. But while a description of early Modern English written in 1909 might well have taken the existence of length for granted, it is important nowadays to re-examine the evidence for such a conclusion.

⁵ *A modern English grammar* 1.79 (Heidelberg, 1928).

⁶ Kökeritz quotes Eugen Zachrisson, *The English pronunciation at Shakespeare's time as taught by William Bullokar* 101 (Uppsala, 1927). Zachrisson has a number of similar spellings, among the more interesting of which is *kyryous*, which strongly suggests the modern pronunciation of *mirror* with stressed /i/.

Of Kökeritz' eight vowels, /æ/, /ʌ/, and /a/ by his own account have only conditioned length, i.e. only before /r/ or /f, s, θ/. If length were phonemic elsewhere, we would be forced to recognize its occurrence even when conditioned, but these examples of conditioned length form no part of the evidence for setting up the phoneme in the first place. Of the remaining five, /i/, /u/, and /o/ have different phonetic qualities when long, so that (again by his own account) length is not here minimally distinctive. An alternate phonemicising might (as in many analyses of Modern English) simply set up the differing qualities as phonemes, and disregard length. Of the remaining two, /o/ has no long-short contrast, since the phoneme occurs only long. The case for phonemic length boils down to one crucial contrast, that which Kökeritz sets up between /e/ and /e:/. If this contrast should prove to be better handled otherwise, so would all other long-short contrasts. Kökeritz discusses this contrast (173-80) under the heading '*a in tale, ai in tail*'. Here he marshals evidence to prove once and for all—Viëtor, Zachrisson, and Ekwall to the contrary notwithstanding—that these sounds had fallen together in Shakespeare's English. All scholars would agree on the facts which he presents, except for a difference in chronology. This difference leads the authorities mentioned above to regard the two sounds as still separate in Shakespeare's speech. The sounds in question were a Middle English 'long *a*' (it is not necessary to consider its exact nature) and a Middle English diphthong spelled *ai* and *ei*, believed by most to have been a single nucleus [æi], though by some thought to have been still two nuclei, [æi] and [ei]. These Middle English sounds fell together. In terms of phonetics, at any rate, the product of the fused sounds is Modern English [ei],⁷ the modern sound having developed a second element exactly identical with the second element of one Middle English source for the modern nucleus. With the exception of Jespersen and Wright all students of early Modern English have assumed that the fusion took place on the basis of a long simple vowel. Jespersen's discussion is still sufficiently cogent to bear quoting in extenso:⁸

It is generally assumed that when *a* and *ai*, *o* and *ou* coalesced, the resulting sound was a monophthong, and that this monophthong was diphthongized in the latter half of the 19th c. But I think it more probable that the coalescence was caused by a diphthongization of the monophthongic sound. The two views may be thus contrasted:

	16th c.	17th, 18th c.	19th c.
<i>ale</i> :	a:l,æ:l	ɛ:l	ɛ:il
<i>ail</i> :	æ:il	ɛ:l	ɛ:il
<i>moan</i> :	mɔ:n	mɔ:n	mɔ:un
<i>mown</i> :	mɔ:un	mɔ:n	mɔ:un
and, on the other hand,			
<i>ale</i> :	a:l,æ:l	ɛ:il	ɛ:il
<i>ail</i> :	æ:il	ɛ:il	ɛ:il
<i>moan</i> :	mɔ:n	mɔ:un	mɔ:un
<i>mown</i> :	mɔ:un	mɔ:un	mɔ:un

⁷ Kökeritz is not very specific as to what he regards as the quality of the modern nucleus, but at any rate he says (164) that in *strange* 'Shakespeare ... had the antecedent of modern [ei]'.

⁸ Jespersen 1.325-6 (originally published Heidelberg, 1909).

Of these two alternatives, the latter would certainly have been universally adopted as involving fewer changes, were it not that the description of the sound given in grammars, etc., of the 18th and of practically two thirds of the 19th c. ignored the diphthong and spoke of all the sounds, in *ail* as well as *ale*, in *mown* as well as *moan*, as monophthongs. The diphthongic character was not generally acknowledged till after the appearance of Sweet's first phonetic works, in the seventies. But the general description of the sounds as monophthongs does not prove much, as this kind of diphthong, with a long first element and a slow upward glide, is not easy to observe or to keep apart from a monophthong. Even such an accomplished phonetician as Ellis failed to observe them in certain cases in his own pronunciation ...

Jespersen's argument is striking, making use as it does of modern phonetics and the criterion of simplicity, so that it is all the more astonishing that it has been more often rejected than accepted. Indeed, the rejection is so surprising that it seems to call for some investigation.⁹ I should suggest that the following beliefs were taken to militate against Jespersen's formulation. First, investigators placed great reliance on the orthoepists. Indeed, Kökeritz (17) still says that they are our most important source of information. But we know now that these writers, like modern school grammarians, are not reliable guides to usage, and are often naively phonemic, presenting contrasts, but not analyzing their nature or their components. Furthermore, the orthoepists of the 17th century often imply a diphthong for *ai*, but none for *a*. Since Kökeritz has conclusively disproved the notion that the two sounds were separate, it is clear that such orthoepists must have been wrong about one sound or the other.

Second, imperfect information about dialects led to the belief that dialect distribution was incompatible with Jespersen's view. Thus Zachrisson observed:¹⁰ 'The chief objection to Jespersen's views that ME *ai* in *day*, and *ou* in *blow* remained diphthongs in early Standard English is that a monophthongic pronunciation of words containing these sounds actually exists not only in many dialects (cf. Wright, EDG. pars. 48, 127), but also in the American pronunciation, which on the whole is based on that of Standard English in the 17th century, and retains many archaic features of pronunciation.' Zachrisson's statement was reasonable enough on the basis of the phonetic data available to him, and so seemed to indicate a late development of diphthongs. But nowadays it is generally ac-

⁹ The original formulation of the theory that *ai* and *ā* both became simple vowels seems to have been in an article by Karl Luick, *Anglia* 14.273-9 (1892). Luick relied on early documents and orthoepists, making no use of dialect materials, and not referring to the phonetic writings of Sweet. The Luick theory was then followed by Viëtor, op.cit. 1.48-52. It was for the first time questioned by Jespersen in the passage quoted. Jespersen was promptly attacked by Eilert Ekwall, *Beiblatt z. Anglia* 21.330 (1910), and by Wilhelm Horn, *Anglia* 35.362-65 (1912). Both these scholars relied on the orthoepists. Thereafter the Luick formulation has been generally followed, as by Zachrisson, *Pronunciation of English vowels 1400-1700* 90 and 222 (Göteborg, 1913), and *Bullokar* 76 fn.; Henry Cecil Wyld, *History of modern colloquial English* (New York, 1920), and *A short history of English* (New York, 1927); Ekwall, *Historische neuenglische Laut- und Formenlehre* 26 (Berlin, 1922); Richard Jordan, *Handbuch der mittellenglischen Grammatik* 241 (Heidelberg, 1925); Luick, *Historische Grammatik der englischen Sprache* 630-5 (Leipzig, 1921 and later).—Jespersen has been followed only by Joseph Wright, *An elementary historical New English grammar* 57 (Oxford, 1924).

¹⁰ *Bullokar* 76 fn.

knowledge that American English normally has [eɪ], even though there are phonemic descriptions which do not regard the diphthongal element as significant. Further, though no one would deny the possibility of long simple vowels in British dialects, it is also generally recognized that the pre-phonemic materials of Ellis and Wright are not an adequate analysis of British folk speech.

The last reason for the continued rejection of Jespersen's interpretation has to do with phonemic analysis. In the several widely current analyses of Modern English which are based on IPA notation, it is usual to write *raid* as /red/ and *red* as /red/, making it easy to forget that phonetically *raid* is [reɪd] or the like. If there is any mixing of phonetic and phonemic levels of analysis, the formulation of long monophthongs in early Modern English, with a phonemic product /e/ in Modern English, seems simple enough.

It would seem then, that there is indeed a possibility that Kökeritz' single minimal contrast in length, /e/ vs. /e:/, could be analyzed otherwise, in terms of a complex nucleus. It is therefore doubly interesting to turn to another section of his discussion, and observe that he makes rather heavy going of certain dialectal developments which would be considerably simpler without the assumption of phonemic length. On page 245 he takes up the fairly common rhymes, found not only in Shakespeare but also in Spenser, which link Middle English 'long u' with Middle English *ou*, as in *allowing* : *growing*. Of these he says: 'Leveling of ME *ū* and *ou* is rare in modern dialects and where it has occurred it must be an eNE development, resulting from the survival of ME *ou* as a diphthong, perhaps [ou], and the diphthongization of ME *ū* to [uɪ] > [ou] rather than to [əʊ] as in the incipient Standard language.' There can be no criticism of the course of development here postulated, except to note that Kökeritz is proposing a dialect in which the Middle English diphthong remained, and so was separate from the Middle English simple vowel. What is interesting is that in the sentence immediately preceding he speaks of the rhyme *down* : *bone* (*Midsummer Night's Dream* 3.2.397-9), which links [əʊ] with an [o:] not from a Middle English diphthong. This rhyme causes him difficulty, and he can explain it only on the basis of /ʌ/ in both words, as in the somewhat doubtful modern /stʌn/ for *stone*, thus postulating an unrecorded pronunciation for both words. The difficulty would not exist were it assumed that early Modern English 'long o' was [ou], since then the explanation quoted above would apply. None of the criticisms I have made of the vocalic system presented are intended to refute Kökeritz' presentation. They are meant, rather, as a suggestion that alternate interpretations are possible, and that the problem is more complex than it seems to appear to Kökeritz. His formulations, though orthodox, now seem somewhat controversial, and should have been given with full discussion.

In the discussion of consonants there is little that calls for more than minor comment. Here, as indeed with the vowels, more use of the materials of the *Linguistic atlas of the United States and Canada* would often have given useful illustrative material. *Philhorse* for *thill-horse* (321), *coted* for *quoted* (330), *Porchmouth* for *Portsmouth*¹¹ (319) can all be illustrated from American speech. Such

¹¹ *Porchmouth* is a widely current form of this place name in Virginia. Kökeritz is inaccurate when he speaks of the development as 'fusion ... of retroflex *r* with *s* to [ʃ]'. It is rather an assimilation of /rts/ to /rɛt/.

variant spellings as *beckles* : *betilles* : *beetles* (305) are referred by Kökeritz to 'the common vacillation between *kl* and *tl* as in *lean* for *clean*'. Not only are such forms as *lean* widely illustrated in the *Atlas* maps, but their explanation is there strongly suggested: neutralization of opposition between /*kl*/ and /*tl*/ after pause. The classic example for medial /*tl*/ > /*kl*/, however, is not *lean* but general American *huckleberry* (British *whortleberry*). The medial /*tl*/ > /*kl*/ probably also involves neutralization of opposition, but the course is different. Medially it is a leveling of two unexploded and simultaneously glottalized stops. In initial position it is rather a direct assimilation of dorsal to apical articulation. *Huckleberry*, with its development /*rtl*/ > /*tl*/ > /*kl*/, suggests one other minor criticism. In his discussion (315-6) of the loss of preconsonantal /*r*/, Kökeritz follows Wyld's statements and examples, which do not clearly distinguish early assimilative loss from the later general loss of preconsonantal /*r*/.

Of the so-called suprasegmental phonemes, Kökeritz makes no attempt to deal with anything but stress.¹² There can be no objection to narrowing the investigation to stress alone, since pitch (terminal or other) would either have to be guessed at from Modern English, or faintly and confusingly inferred from punctuation. But Kökeritz works with a framework of three stresses only, which is inadequate for a language where four stresses are significant. Even these three stresses, moreover, are considerably confused in transcription. Thus he says (344), 'subtle gradations of sentence stress have not been attempted'. This means in practice that he writes secondary stress only in polysyllabic words. Monosyllabic words are written only with a strong or a weak stress.

Kökeritz' conclusions on word stress have necessarily been reached on considerations of metre, and here his understanding seems to me to leave a good deal to be desired. Thus he says (333): 'Verse stress is not the same thing as word stress. Generally the two coincide ... Nevertheless, infinite variations in rhythm may be set up by creating a tension between the verse stress on the one hand and the word accent on the other.' I should be the last to deny that natural stress and verse stress are different, or to deny that there are tensions between them. But a statement which does not define more exactly how this tension works—even though this is the usual critics' formulation—is not precise enough to be an instrument for reaching scientific conclusions. The problem, of course, is a very real one, and I do not know of any account of English versification which resolves it satisfactorily. I should propose the following solution. English verse stress recognizes only a strong and a weak. This means that the two extremes of natural stress (primary and unstressed) are fixed in verse. The two middle grades of natural stress (secondary and tertiary) are variable in verse, according to

¹² Kökeritz has a curious statement (3): '[We must] take into account not only the basic elements of his speech, that is, vowels, consonants, and stress, but also such secondary characteristics as pitch and resonance of voice and the principal patterns of intonation and rhythm. Most of these secondary features are extremely difficult to analyze fully and to describe intelligibly when we study the speech habits of our own generation even with the aid of modern recording machines.' — Work by Wells, Pike, and Trager (to name only a few), has shown that pitch is a part of the phonemic system, not on the same level with resonance, and there are now at least two reasonably accurate systems for analyzing and describing pitch intelligibly.

whether they are adjacent to stresses stronger or weaker than themselves. The practice of poets would seem to be to keep their variations within this limitation, or run the risk of being condemned by readers and critics as unskillful.

If such a formulation should prove to be well founded, it would very considerably affect conclusions about Shakespeare's word-stress which are based on verse patterns. Thus Kökeritz (392) gives *absurd* as generally stressed on the second syllable but in one occurrence on the first. He discusses the line on which this conclusion is based as follows (334): 'The interdependence of verse stress and word accent is demonstrable in almost all the instances just quoted; cf. e.g. *lick absurd pomp*.' The full line is: *No, let the candied tongue lick absurd pomp* (*Hamlet* 3.2.65). From the partial scansion which Kökeritz gives, it would seem that he is scanning with absolutely regular iambs:

No, lét | the cán | died tóngue | lick áb | surd pómp

But if, on the other hand, the line is read with a pattern which is natural in Modern English (not necessarily the only natural pattern, of course), the result is:

Nó, lèt | thě cán | diéd tóngue | lick ăb | sùrd pómp

The metrical irregularity of the line is primarily in the occurrence of trochees in the first and fourth feet, a liberty which many poets have both practiced and valued. There is further irregularity only in the fact that two tertiary stresses count as poetic weaks. I should therefore regard as unproved the conclusion that *absurd* could have initial stress. And as a result, it seems to follow that all of Kökeritz' conclusions about word-stress need to be re-examined. Nor is the untrustworthiness of method in investigating stress of importance to historical phonologists alone. His failure to recognize the complexity of the relations between natural and verse stress would appear to be at least partly responsible for his view of Shakespeare as a regular, syllable-counting versifier.

In his treatment of historical developments, as might be expected of a scholar who has done distinguished work in dialectology, Kökeritz follows Wyld in his explanation of one of the knottier problems of development, the coalescence of Middle English open and close 'long e'. This he presents not as a complex series of sound changes within the Standard language, but as the victory of one dialect type over another.¹³ Kökeritz also considerably improves Wyld's chronology. Wyld held that the two vowels remained separate throughout the Middle English period, and that only in early Modern English did there develop a dialect type in which the two fell together. Kökeritz pushes the dialectal coalescence back

¹³ Wyld, *Colloquial English* 210: 'We must assume, therefore, that the [ē] from earlier [e] was differentiated among different classes of speakers—whether in a Regional or a Class dialect I am unable at present to say—into two types, one of which retained the old [ē], while the other gradually raised this to [i].' This statement, which accounts for the variant rhymes of poets, the contradictory statements of the grammarians, and known variants in old-fashioned speech such as *speak* with the nucleus of *cake*, contrasts with the relatively complex explanation given by Luick, *Historische Grammatik* 597, which has the disadvantage, among others, of requiring the assumption of a difference (not needed for anything else) between a long open and a long close *i*.

into Middle English, and suggests the way in which it occurred. He points out (196) that Kentish Middle English had inherited only the close vowel from Old English, and that in the eastern dialects the open vowel became close 'before dentals'. In other words, in eastern and southeastern dialects there was a considerable reduction of opposition between the two. There is only one minor difficulty in his formulation. He names Middle English /r/ among the dentals which raised open to close 'long e'. It is possible that the two vowels were reduced before /r/, but if so, it must have been on the basis of the open rather than the close vowel, since the tendency of the Middle English trill was to lower a preceding vowel. I have long felt that this was what happened, and that this lowering is one of the reasons why the coalescence must have taken place early. The replacing of the product of an open vowel with the product of a close vowel, as in *cheer* for *chair*, can only have occurred late, after the /r/ had ceased to be a trill. Another of Kökeritz' most valuable contributions also occurs in the discussion of these vowels before /r/. He gives a table of rhymes before that consonant, which shows clearly that variation existed, but that rhymes which agree with modern distribution of the vowels are the most frequent type. That is, modern usage was already on the way to being established.

The final sections of this review will be concerned with the assumptions under which Kökeritz handles his material. I shall take up evidence from spelling, puns, and rhyme, and conclude with a brief statement on textual problems. Kökeritz handles spelling evidence excellently. His attitude is that occasional spellings can only be interpreted in the light of a norm of the speller's dialect and social status; it is the relation between the departure and the norm, rather than the spelling itself, which is significant. The observation is sound, though I could wish that it had led to a systematic presentation of graphemics. But that is asking a good deal, since there is no published account of the graphemics of Shakespeare or anyone else. What Kökeritz does instead is of real value; he makes wise and extensive use of spellings from the same time and place to throw light on those in Shakespeare, and he also attempts as far as possible to separate the printer's spellings from those which he believes to be characteristic of the manuscript.

His handling of puns is somewhat less satisfactory. He recognizes two types of word play: those with a difference, which he calls jingles, and those involving absolute identity, which he calls homonymic puns. He gives us, however, no criterion for judging whether any particular word play is homonymic (and so of phonological importance) or a mere jingle. I know of no way in which the necessary distinction could be set up without an extensive study of English punning habits, and perhaps even then the result would be negative. It may, however, be worthwhile to make some tentative suggestions about the differences permissible between the like entities in a modern word play. Word plays disregard supra-segmental phonemes rather consistently, as in *Did you ever see a horse fly?* A pun like *bare skin : bearskin* needs at least near identity in the segmental phonemes, whereas the one ascribed to the playwright George Kaufman, *One man's Mede is another man's Persian*, does not. The principle seems to be that as the number of identical phonemes increases, the need for identity in any corresponding pair

is lessened. Finally, word plays apparently demand a like syllabic structure, independently of the number of like phonemes.

Tentative as these suggestions are, they can be brought to bear on some of Kökeritz' discussions. He cites (67) a jingle ascribed to Dr. Tobie Matthew: *Our Paradice is a paire of dice, o[u]r almes deeds are turned into all-misdeeds ...* Kökeritz would seem to be right in supposing that this play is evidence that Matthew used a dissyllabic dialectal pronunciation of *almes*. Such a conclusion makes use of likeness of syllabic structure. But his inference from the preceding play that Matthew used /æ/ in *paradise* is much less secure; this would be a good modern pun even with some totally unpredictable vowel like /o/. In short, phonological evidence from puns is dangerous unless some means can be devised to keep it from cutting both ways. I have, for instance, heard punning habits used to prove both identity and difference in the line from *The Merchant of Venice* (4.1.123), *Not on thy soale: but on thy soul harsh Jew*.

In handling rhymes, Kökeritz has done a service in pointing out a number of rhymes which would not now be regarded as perfect. These consist of pairs like *sign : time*, *is : amiss*, *beseech : liege*, *halt : talk*. That is, Shakespeare rhymed consonants of roughly the same order or manner, even though there was a phonemic difference between them. It is interesting that this convention of rhyming survives in contemporary folk poetry, but does not occur in the poems of men like Chaucer, who were writing in a learned rather than a popular tradition. These Shakespearian rhyming habits, since they agree in one characteristic with the habits of popular poetry, make it less likely that Shakespeare preserved syllabic regularity, as Pope did. He is more likely to have followed the pattern, in both rhyme and syllabic structure, of such popular pieces as the proverb,

*A stitch in time
Saves nine.*

While Kökeritz is both sound and interesting in his discussion of consonant rhymes, he seems to me more questionable in his treatment of rhyming vowels. Several times (e.g. 214) he speaks of rhymes of different but similar vowels. This is a matter of great importance for historical phonology; it is, indeed, the very point on which Kökeritz has condemned Viëtor. No one would claim that all Shakespeare's rhymes are perfect; the question is rather, given an imperfect rhyme, is it evidence of similarity? It would seem to me a safer assumption, when a rhyme cannot be explained on the basis of dialect or tradition, that it has no evidential value.

Still another unsettled question in Kökeritz' treatment of rhymes is how many phonemes must be identical. The rhyme *confusion : division* (211) puzzles him, but he remarks that the rhyme *posteritie : obscuritie* 'may be a rhyme in -ty'. If one pair can rhyme on the basis of one syllable alone, why not the other? Again (132, 232, 320), he discusses a pun on *nothing : noting* and a rhyme *nothing : a-doting*. In support of identity for the two words he cites a dialectal sound-change of final /θ/ > /t/, which apparently underlies Shakespearian variants such as *mote* for *moth*. A pronunciation such as [ˈnotɪn] for *nothing* is then possible, but neither the puns, which need not be identical, nor the rhyme, which

may be based on only the concluding phonemes, or involve the inexact type of consonant rhyme discussed earlier, prove identity. The phonological discussion has a bearing on the title *Much Ado About Nothing*, which, it has been suggested, contains a pun on *noting*. Kökeritz accepts the pun, but redefines *noting* as 'brand with disgrace', rejecting the usually suggested 'evesdropping'. The evidence for the pun seems shaky, since it is not necessary to assume identity of pronunciation, and in the absence of identity one wonders how a title, given as it ordinarily is without context, could have made the pun evident.

As with puns, the value of Kökeritz' evidence would have been greatly increased by a systematic survey of Shakespeare's rhyming conventions, made by working outward from the unquestioned and perfect rhymes. It would undoubtedly have told us much, not only about the value of rhyming evidence, but about the nature of English literary conventions. Study of that sort, indeed, is one of the ways in which linguists and literary students can work together for the total understanding of literature.

The final matter is Kökeritz' attitude towards his text. It is interesting that analytical bibliography has been experiencing a vigorous growth in the last two decades, which quite closely parallels the growth in linguistics. Bibliographers have developed ways of weighing variants much less dependent on subjective judgment than those used by older or less exacting editors, and in consequence have abandoned more and more brilliant emendations in favor of the reading of the text. In an ideal scholarly world, the two developing sciences would go hand in hand, since the bibliographers, who handle linguistically structured material, obviously need the results of linguistic examination of such materials, and since the linguists, particularly historical phonologists and lexicographers, need to know the methods of modern bibliography if they are not to expose themselves by reckless emendation, or to be at the mercy of unreliable editors. An example of rather radical treatment of the text can be found in Kökeritz' suggestion (178), '*Haile* (*Macbeth* 1.2.5) is usually taken to be disyllabic, but I can see no reason for this phonological dodge; the missing half-foot (thesis) can easily be supplied by a pause or by a second *haile*, which may inadvertently have been omitted by the printer.' This is to avoid a phonological dodge by the alternative of a dodge either in metrics or text.

A more important textual matter is the brief discussion of the crux *too, too solid flesh*. In this Kökeritz is following John Dover Wilson,¹⁴ probably at one remove. What is involved is a recurrent form *sallied*, which, as usual, Kökeritz emends to *sullied* (242 and fn.). In the *Hamlet* line, the quartos read *sallied*, and only the folio has the usually accepted reading *solid*. Wilson emends to *sullied* and then eloquently defends this against the usual *solid*. But *sallied*, or a derived form of it, occurs twice in *Hamlet*, once in *Love's Labour's Lost*, and once in *Patient Grissell* by Dekker and Chettle. In all these occurrences, 'soiled, sullied' would fit the context, though the meaning 'assaulted', which has been suggested, would not. Such regular recurrence would seem to be fairly good evidence that an Elizabethan *sallied*, with the same meaning as *sullied*, actually existed, and that the *Hamlet* alternatives are this form and *solid*, not *sullied* and *solid*.

¹⁴ *The manuscript of Shakespeare's Hamlet* 2.307-16 (New York, 1934).

The form is at least possible as a borrowing from French *sale*. Editors who use one of the four emendations to support another, and even build up an inference of confusion between *u* and *a* in Shakespeare's manuscript, are being somewhat circular. Another even more minor difficulty is that Kökeritz follows, somewhat tentatively, the alternate suggestion quoted by Wilson, that *sallied* in *Hamlet* is a mere spelling of *solid*, with *a* for unround *o* as elsewhere in Shakespeare. The suggestion does not seem to be a very good one, since not only would it not account for the other occurrences of *sallied*, but a quick check of forms in the NED does not make it seem likely that a syllable /-ɪd/, where the /-d/ is not a grammatical ending, would have been spelled *-ied*.

It is not my purpose to argue the virtue of *sallied* as against *solid*, but rather to point out that in following a radical editor, Kökeritz has been led into a somewhat unlikely dialectal suggestion to explain a spelling, and has missed a word which is of some possible importance to Elizabethan lexicography. The whole textual crux is in itself unimportant enough. But none the less it illustrates a basic quality of Kökeritz' book. It is characterized throughout by enormous labor, by the collection of large amounts of new and important evidence, and by new and important results. Yet a book with a more explicitly stated and rigidly controlled method would have been even more valuable. The reader can be grateful for the great services Kökeritz has performed, at the same time that he laments the absence of others still greater.

The triumph of the English language: A survey of opinions concerning the vernacular from the introduction of printing to the Restoration. By RICHARD FOSTER JONES. Pp. xii, 340. Stanford, Cal.: Stanford University Press, 1953.

Reviewed by ARCHIBALD A. HILL, *Institute of Languages and Linguistics*

Rightly or wrongly, linguists expect an attitude of ethno- and glotto-centrism to be the normal view of the native idiom in any community in which culture and language are still going concerns. Only when a language comes to be characteristic of a disintegrating culture, or of a depressed minority, do we expect the native speakers of a tongue to assume that it is an inferior way of speaking—a barbarous jargon. It is therefore of real value for understanding how people have thought about language, to have the history of native attitudes towards English told over again, and in detail. Baugh has made us all familiar with the successful Renaissance attempts to 'enrich' the language by borrowing, and the gradual growth of necessarily less successful attempts to regulate and fix the language at what was believed to be its apogee. Jones gives us this and much more. He begins by pointing out that the normal 15th-century attitude was that English was adequate to express bare meaning, but lacked polish and eloquence in contrast not merely with the ancient languages but with contemporary languages as well. He tells us that later the vocabulary was felt to be inadequate for propagating the new learning, and that it was therefore a scholar's patriotic duty to translate, and in so doing to enrich. He shows that the tremendous literary flowering of the Elizabethan period brought with it a revolution in attitudes toward English. since now for the first time writers began to speak of the lan-

guage as one of literary sophistication. Finally he traces the revolt of the Puritans against the educational dominance of Latin, a language regarded as medieval, aristocratic, and useless in the practical sphere. In this, as in much else, the Puritans foreshadow the 20th century.

In describing these successive attitudes, Jones brings to bear a wealth of learning: one has the impression that almost everyone who expressed himself in print on the vernacular, between 1475 and 1650, is honored with a quotation. More important, he relates these ideas to the wider world, showing for instance, how Protestantism led to a demand for books of instruction—first in theology, later in other subjects—for the benefit of the common man, and how the emergent linguistic nationalism of the 17th century was based on the grotesque veneration for Germanic languages of the continental writer Goropius, who seriously proposed that German was the language of the Garden of Eden.

When a scholar has written a book with a broad point of view, it is perhaps unfair to demand that the point of view be still broader; yet there is at least one attitude which demands explanation but does not receive it. This is the attitude of the 15th century that English is inferior in literature and scholarship. Jones seems to feel that such an attitude was normal at the end of the Middle Ages, and does not comment on it. Baugh has explained this linguistic inferiority complex on the ground of the subjection of English to French after the Norman conquest, but one wonders whether this is the whole story. Part of it may well go back to the Christianizing of England, when the English were brought into an alien cultural circle, and made to feel that they were at the periphery of it, not at the center. For centuries thereafter the heart of religious activity, of learning, even of commerce and cuisine, pulsed from the continent. Englishmen came to think themselves inferior, even as early 19th-century Americans felt themselves gauche in the presence of British lecturers. Yet to this reviewer it still seems strange that the vernaculars should have been despised so long after the emergence of the new states. This paradox in the history of ideas indicates that our understanding is still incomplete.

In certain ways Jones could have profited by more knowledge of the discipline and even the speculations of linguistics. For instance, it is not clear that early complaints about the English language are really concerned with the absence of traditional literary style and ornament, and touch the language proper only in the relatively marginal area of learned vocabulary. The reader misses the necessary qualification that the medieval English vocabulary cannot be proved to have been smaller than the vocabulary after the addition of the hosts of ink-horn terms. Addition of specialized items is not the same thing as increase of vocabulary, or at least cannot be assumed to be the same thing in the present state of linguistics, when all that we know about the size of vocabulary is that we cannot measure it.

A second complaint is that the discussion of translations and translators does not make clear what was happening, and that what was happening, as linguists have pointed out, was a general European process. What was going on was a sort of pooling of an accepted body of European writing: translation back and forth from the classic tongues to the vernaculars, from the vernaculars to the classics,

from vernacular to vernacular, until there developed that mutual inter-translatability which is one of the most striking characteristics of the European community. Striking as it is, however, it is a feature that one does not appreciate until one tries to translate from a language sharing the quality (say English) to one which does not (say Cherokee). If this interpretation of the process is right, then the translators in the Renaissance were not merely enriching the English language, but contributing to a European linguistic and intellectual unity which is very real, in spite of the diversity which conceals it.

Jones expresses himself in his preface as distrusting the tendency of scholarship to reduce data to patterns whose neatness and symmetry, he feels, are no guarantees of their trustworthiness. This is a common view among humanistic scholars, and in this book is responsible for the inclusion of large numbers of quotations, many of them repetitious. A structural linguist would agree that in so far as symmetry and economy are achieved at the expense of completeness, they are indeed untrustworthy; but he can be pardoned for wishing that there had been more analysis and less piling up of testimony. Paradoxically enough, there is some simplification which seems doubtful. Jones presents his various attitudes towards the vernacular as roughly chronological. Yet when one observes the recurrence of quotations, now in one chapter, now in another, it is evident that there is more chronological overlap than his chapter headings allow for. Yet these criticisms are surely minor, and Jones has written what will remain the definitive study of this particular subject on the edge of linguistics, within the time limits which he has set himself. The only further demand that seems legitimate is that he, or another, extend the study to earlier and later periods.

On the history of some problems of English grammar before 1800. By IVAN POLDAU. (Prague studies in English, Vol. 7; *Práce z vědeckých ústavů*, No. 55.) Pp. 322. Prague: Philosophical Faculty of the Caroline University, 1948.

Reviewed by PAUL L. GARVIN, *Institute of Languages and Linguistics*

Poldauf's study of English grammatical thought between 1586 and 1800, based primarily on materials in the Bodleian Library at Oxford, fills an important lacuna in the history of linguistic theory. Most histories of linguistics I have seen (Vilhelm Thomsen, Holger Pedersen, Tadeusz Milewski) treat the grammatical ideas of the ancients in some detail, then devote most of their space to developments since Rask and Bopp, glossing over the intermediate period by some convenient reference to scholasticism and the persistence of the traditions of antiquity. For those who wish to know more about this supposedly dark age of linguistics, Poldauf's book presents a wealth of general European material in addition to his major interest, showing the gradual increase of interest, originality, and acumen.

The book consists of two major sections—Introduction: English grammars before 1800 (45–150), Some problems of English grammar (151–311). It contains two bibliographies: one of books and articles consulted (17–20), and one of Eng-

lish grammars before 1800 (21-39); it has also an index of persons (313-20) and one of subjects (321-2), making reference very convenient.

The first half of Poldauf's study consists of an annotated bibliography of English grammars published between 1586 and 1799 and available in the Bodleian library. The list is extensive, and includes some items not cited in earlier bibliographies, such as Kennedy's. The chronological arrangement has the advantage of making clear the sharp rise in interest in English grammar in the second half of the 18th century, when the Campbells and the Lowths were prescribing. The bulk of Poldauf's entries fall between 1750 and 1799. Of special interest are the references to English grammars written on the continent.

The chronological order makes it difficult, of course, to pick out the more significant grammars and to trace the history of ideas. Some of the works discussed are concerned rather with spelling and pronunciation than with grammar. However, Poldauf gives due weight to the influence of Swift and Johnson, to the notion of an academy, and to the development of the prescriptive tendency. His work is a useful adjunct to the special studies of Funke and S. A. Leonard. Poldauf's annotations should prove convenient to students who are working with particular topics in the field of early modern English grammar.¹

In his second section, Poldauf discusses the treatment by the early English grammarians of several specific problems of English grammar, to each of which he devotes a chapter: Parts of speech (151-82), Articles (183-98), Gender (199-210), Number (211-39), Adjectives and comparison of adjectives (240-62), Tense (263-311). Each chapter deals with the treatment of the problem in historical sequence, integrated into the history of general European thought on the subject, and evaluated in terms of the adequacy of the handling of the material as judged by modern standards. The author points to many instances of keen insight into the synchronic functioning of English, in a period when most historical and etymological interpretations of language were still rather wild. Certain of Poldauf's chapters are of considerable interest from the standpoint of modern English grammar, since his running commentary contains many original interpretations of English structure; note his astute analysis of countability as a factor in the semantic content of the number category, or his equally stimulating discussion of gender and its relationship to sex. In general, the interest of this section of the book far exceeds that of a history of early English grammatical thought.

The phonemes of English: A phonemic study of the vowels and consonants of standard English. By A. COHEN. Pp. [viii], 127. The Hague: Martinus Nijhoff, 1952.

Reviewed by GEORGE L. TRAGER, *Foreign Service Institute*

This book is a doctor's thesis prepared under the supervision of A. J. B. N. Reichling, and thus presumably at the University of Amsterdam (identified

¹ For the information and the critical opinions in this paragraph I am indebted to Rudolph C. Bambas, my former colleague at the University of Oklahoma.

in the Preface only as 'this University'). The author studied at the University of London, and his work is largely in the tradition of Daniel Jones and his followers, though with a large admixture of that continental European approach which the reviewer feels to be philosophical rather than scientific.

The book is organized as follows: Chapter I. Introduction (1-13); II. Problems connected with phonemic analysis: 1. Fundamental points (14-8), 2. Significant functions of sounds (18-26), 3. Phonemes and their interrelations (27-36), 4. The positive factor in the habitus of the phoneme (36-7), 5. Consonant and vowel phonemes (37-41); III. The consonant phonemes of English: A. Inventory (42-51), B. Classification (52-74); IV. The vowel phonemes of English (75-107). An Appendix (108-27) contains a list of publications consulted, and indexes of authors, words, and subjects.

The first third of the book (Chapters I and II) contains the setting forth and justification of theory and method without which a dissertation in this field is considered incomplete. The reviewer finds little new in these discussions; they are the old arguments—chiefly philosophy, little science.

The English considered is only 'Standard British English' (1), though American works are held to be 'of great importance' (3). The phonemes presented are only vowels and consonants (19).

The plosives are /p b t d k g/ (Cohen uses slant lines and brackets in the now almost universally accepted way); the affricates [tʃ] and [dʒ] are considered to be clusters; the nasals and laterals are /m n ŋ l/, of which /ŋ/ is discussed at length; the fricatives are /f v θ ð s z ʃ ʒ r h/, of which the last again is discussed at length; the semivowels are /j w/. The sound [ç] is interpreted as /hj/, though the wording is not clear (58). As for [ʌ], the reviewer simply doesn't understand the remark, 'as no significant use is made consistently between [hw] or [ʌ] and [w] there is no need for us to accept a cluster /hw/' (58); footnote 35 at this point observes, 'Yet Bloch and Trager ... arrive at the conclusion that /hw/ ... should be interpreted as a biphonemic cluster in American English.' The reviewer may comment that in spite of Cohen's 'Yet', the cluster /hw/ seems inescapable in any kind of English that has something like [hw] or [ʌ].

The 'short' vowel phonemes are established by the contrasts between [i e æ ʌ ɔ u] (75). It is concluded that [ə] is a separate phoneme (77), and that [u] is really separate from [ʌ] (78). The long vowels are [i: ɑ: ɔ: u: ə:] (78); the diphthongs are [ei ou ai au oi iə eə əə uə]. Of these all except the last four are held to be unit phonemes (99). The 'centring' diphthongs (except [əə]) are interpreted as 'contextual variants of the vowels /i/, /u/, and /æ/' (101), and [əə] is identified with [ɔ:].

The reviewer holds that much of the theory that Cohen follows is inapplicable or misleading, that the phonetic data are presented inadequately at many points, and that the conclusions are erroneous as regards [tʃ dʒ], [ə], and all the 'long vowels' and 'diphthongs'. His views are completely set forth in his and Smith's *Outline of English structure* (SIL, Occasional Paper No. 3, 1951). A comparison of the phoneme lists may be given for the reader's convenience. (T-S stands for Trager and Smith; an asterisk marks unit phonemes represented by digraphs.)

COHEN	T-S	COHEN	T-S	COHEN	T-S	COHEN	T-S
p	p	v	v	i	i	ei*	ey
t	t	ð	ð	e	e	ou*	ow
k	k	z	z	æ	æ	ai*	ay
b	b	ʒ	ʒ	u	u	au*	aw
d	d	m	m	ɔ	ɔ	oi	oy
g	g	n	n	ʌ	ə	[iə]	ih
tʃ	č	ŋ	ŋ	ə	ə, i	[eə]	eh, əh
dʒ	j	l	l	i:	iy	[ɔə]	oh, əh
f	f	r	r	a:	ah	[uə]	uh
θ	θ	w	w	ɔ:	oh, əh	?	o
s	s	j	y	u:	uw	?	a
ʃ	š	h	h	ə:	əh, ih		

Kleinasiatiscche Nachträge. By JOH. SUNDWALL. (Studia Orientalia edidit Societas Orientalis Fennica, Vol. 16, No. 1.) Pp. 50. Helsinki, 1950.

Recueil d'onomastique hittite. By EMMANUEL LAROCHE. (Ouvrage publié avec le concours du Centre National de Recherche Scientifique.) Pp. 153. Paris: Librairie C. Klincksieck, 1952.

Reviewed by ALBRECHT GOETZE, *Yale University*

Sundwall's *Kleinasiatiscche Nachträge* are meant as a supplement to his book *Die einheimischen Namen der Lykier nebst einem Verzeichnis kleinasiatischer Namenstämme* (*Klio*, Beiheft 11; 1913). When the latter book was compiled, Anatolian studies were in their infancy. Sundwall followed the general assumption first posited by P. Kretschmer in his *Einleitung in die Geschichte der griechischen Sprache* (1896), that the peoples of western and southern Asia Minor¹ are related to one another and represent a family of languages which is neither Indo-European nor Semitic. The prominence given to the Lycians resulted from the circumstance that Lycian was then the only language of Asia Minor for which epichoric material² was available.

In the meantime the Hittites have been rediscovered; the clay tablets excavated at Boğazköy, presenting copious texts in Hittite and less copious texts in other ancient languages, have pushed back Anatolian history to the beginning of the 2nd millennium and put Anatolian research on an entirely new basis.³ Moreover, the 20th century B.C., the century preceding the creation of a Hittite state, has been illuminated by thousands of 'Cappadocian' tablets written in Old Assyrian by merchants from Assyria; the vast majority were found at a site near the 'Kültepe' which covers the ancient city of Kaniš.⁴ Finally, inscriptions written with so-called Hittite hieroglyphs, used in Asia Minor and in northern Syria for inscriptions on rock and on buildings from ca. 1600 B.C. down

¹ This qualification, lacking in Kretschmer's statement, should be added.

² E. Kalinka, *Tituli Lyciae lingua Lycia conscripti* (*Tituli Asiae Minoris* I, 1901).

³ A survey of the results is presented in A. Goetze, *Kleinasien* (*Kulturgeschichte des Alten Orients*, Handbuch der Altertumswissenschaft 31, 1933).

⁴ Op.cit. 61 ff.

to about 700 B.C., are gradually becoming intelligible and offer us a new language and fresh historical information.⁵

Against Kretschmer's original assumption it is today clear that 'Hittite' is after all related to the Indo-European languages; the texts themselves call that language 'Neshite'. More than that: two other languages attested at Boğazköy, Luwian⁶ and Palaic,⁷ have been shown to be closely related to Hittite; they form together with it an 'Anatolian' group of languages. Another member of this group is the language concealed behind the hieroglyphs. The older population of Asia Minor is represented by the language which the Boğazköy texts call Hattic.

In this situation the names of native Anatolians contained in the Old Assyrian texts from Kültepe become increasingly significant. The question as to their appurtenance, wholly or in part, to the Anatolian group or to Hattic, demands investigation. As will be shown elsewhere, the Kültepe texts testify to the presence in Anatolia of one or the other of the Anatolian languages as early as the 20th century B.C.

Since Sundwall's first book the linguistic problems of early Asia Minor have thus changed completely. They have acquired much larger scope as well as much greater depth. The chief problem posed by the Anatolian proper names contained in the inscriptions of the Hellenistic and Roman periods with which Sundwall dealt must now be formulated thus: Is the language (or languages) in which they originate traceable in the earlier material?

Nowadays the 'Lycian' names play a much smaller role than they did forty years ago. Nevertheless, the collection of additional names from Anatolian inscriptions of the Hellenistic and Roman age which Sundwall presents in his *Nachträge* remains valuable. But the comparison, coordination and correlation with the names of older age is not advanced by the very eclectic inclusion of a few Boğazköy names. Here Laroche's *Recueil* comes in. With its approximately 1000 names listed from Boğazköy texts, it circumscribes the magnitude of the available treasure of names. Laroche, in turn, does not exhaust the names of native Anatolians contained in the 'Cappadocian' tablets,⁸ and of the names of the hieroglyphic inscriptions he can, in the circumstances, merely give the relatively few of which the reading is reasonably certain.

The linguistic analysis of this large material has hardly begun. With an increasing knowledge of the respective languages much refinement will be possible in the future. But the main result, it seems, can already be outlined: there is no doubt a large element of continuity which can be traced from our oldest sources

⁵ The decipherment (mainly due to H. T. Bossert, E. Forrer, I. J. Gelb, B. Hrozný, and P. Meriggi) has been put on a wider and safer basis by Bossert's discovery at Karatepe of a bilingual text in hieroglyphic Hittite and Phoenician. For orientation compare M. J. Melink, *Bibl. Orient.* 5.141-50 (1950); H. G. Güterbock, *Eranos* 47.93-115 (1949); I. J. Gelb, *Bibl. Orient.* 5.129-41 (1950).

⁶ Additional work on Luwian is impending.

⁷ Otten, *ZA NF* 14.119-45 (1942).

⁸ The list contained in F. J. Stephens, *Personal names in Cappadocia* (YOS Res 13:1, 1928), is antiquated by the publication of many additional texts. For the Anatolian names compare now E. Bilgiç, *AfO* 15.2 ff. (1952).

right down to the Hellenistic age. Despite Hellenization and Iranization, the epichoric languages still spoken in the first centuries of the Christian era in certain areas of Anatolia, are, at least in part, offshoots of those languages which emerge around 2000 B.C. in our earliest records.

A few details may be selected for critical remarks.

In his earlier book Sundwall had reduced the names he listed to 'Lycian' forms. This practice has now been abandoned and a purely alphabetic arrangement substituted. The change of procedure can only be approved. Lycian is only one of the younger Anatolian languages and not too well known itself. If it was odd in 1913 to reconstruct (often quite arbitrary) 'Lycian' forms for names from Cilicia, Pisidia, and Lycaonia, it would be an impossibility today. It is not always easy to recognize 'Anatolian' names and some material has no doubt crept in that is not pertinent. Much as the specialist in Anatolian epigraphy may find to correct, scholars dealing with the earlier languages of Asia Minor will nevertheless have to use Sundwall's lists and they will be grateful for what they will find in them.

Laroche for his part not only offers lists of the proper names of the Boğazköy texts (and related material), he also contributes notably to their analysis and interpretation. The entries are given in a 'broad transcription'; one would have preferred a sign-by-sign transcription according to the spelling of the original texts. The spelling and its variants may have significance for the interpretation. Hittitologists will take special notice of the discussion of readings proposed for theophorous and other elements; note in particular the new and convincing argument for the equation $^{\text{D}}\text{KAL} = ^{\text{D}}\text{Inar(aš)}$, 78 ff.

Laroche comes to Hittite studies from the classics and from comparative philology. In matters Akkadian he is more dependent on other scholars. Thus he takes over from Friedrich (*Arch. Or.* 17:1.251) the non-existent female name $^{*}\text{Ki-el}$; read instead $^{\text{SAL}}\text{KI.SIKIL}$ (or KISKIL), which is the Sumerian word for 'maiden' (Deimel, *ŠL* 461.271).⁹ The alleged Cappadocian name *Luarrah/kšu* utilized on p. 105 for proving interchange between *h* and *g* originates with Stephens (*Personal names of Cappadocia* 55), the hyphens having been removed. In Stephens' quotation *ak* is simply a misprint for *ah*; moreover the name reads everywhere (add KTHahn 28a 4, b 6 ff.) *Lu-úh-ra-ah-šu*.

ZH guide: An introduction to Sinology. By GEORGE A. KENNEDY. (Yale University Sinological Seminar.) Pp. [viii], 171, offset. New Haven: Yale University, Far Eastern Publications, 1953.

Reviewed by EDWARD H. SCHAFER, *University of California*

As Kennedy announces in his Foreword, this volume has little in common with so-called 'practical' language studies, but it is, for all that, an eminently practical book. Kennedy rejects the doctrine, which is still widely accepted, that knowl-

⁹ It seems time to correct this mistake, so that its perpetuation may be avoided. In the mean time it has been taken over also by T. H. Gaster, *Bibl. Orient.* 9.85 (1952). The Hittite reading is unknown; one may perhaps think of $^{\text{SAL}}\text{dammara}$ - (*KUB* V 6 II 16, 36, 46; XVI 16 obv. 13, 18, 23, 26, rev. 19; XVIII 3 7).

edge of some modern Chinese vernacular is an important prerequisite for the study of the literary language—as if the serious student of Latin ought first to become conversant with the modern dialect of Tuscany. He assumes, furthermore, the rather old-fashioned principle that historical depth in scholarship is valuable for the student of contemporary affairs. This notion is given some lip-service nowadays, but in practice it is not greatly honored. Kennedy offers, in short, an elementary textbook for the student of sinology which deals with philological matters and nothing else, and treats them with the vigor, precision, wit, and charm which they deserve but rarely get. He discusses serious problems of the utmost difficulty with honesty, and neglects the time-honored trivia which have so often engaged the attention of students of the Far East who have lacked the benefits of classical training, linguistic scholarship, or natural discernment. The student who makes use of this book must be prepared hereafter to swallow the clichés of all factions of linguists and philologists with large doses of salt, and to see problems of analysis and exegesis for what they are in their proper contexts. Kennedy has not solved every problem which may confront the beginning student, but he has, by and large, shown what the essential problems are, and what procedures offer hope for their solution. Let us take a closer look at the book itself. The ten chapters will here be discussed one by one.

1. DEFINITIONS. Here the author announces the nature and uses of the Chinese reference work to which his own book is a guide (though 'guide' is a modest title), namely the *Tz'u hai*, or in his romanization ZI HOJ, whence the abbreviation ZH which appears in the title. He discusses the general plan of this useful encyclopedic dictionary, and explains his own terminology, in particular a set of alphabetic symbols which he employs to represent the syntactical predilections of the morphemes contained in the dictionary. For example, 'N = noun, not capable of taking the negative prefix', 'j = adjunct to a noun', 'Nv = a noun functioning in context adverbially'. Some such classification of the elements of the Chinese lexicon has long been desirable, but has never heretofore been made available by the lexicographers of the literary language. Your reviewer is a little uncertain, however, as to the method whereby one ascertains that the word 'XWAY 125', for instance, is 'basically' a verb 'draw, paint' and only secondarily a verb functioning as a noun 'painting, picture'. The opposite statement might be equally true; but if not, Kennedy has not shown here why not. It would seem clearer to say: 'V, translatable in context by English noun-form', 'Vn, translatable in context by English verb-form'. The word classes of literary Chinese are not readily definable, and the categories 'noun' and 'verb' are at best rough and ready approximations based on relative frequency of occurrence in the several possible positions in a sentence. Later in this review I shall indicate one way in which this problem might be handled. However, Kennedy states, 'For the theory and analysis of classical syntax one should consult A Grammar of Mencius, published in this same series.' No doubt the reader would find it advisable to do so.

Though admittedly far from fluent in the usages of descriptive linguistics, your reviewer finds it difficult to understand the sense in which Kennedy employs the expression 'close juncture' with reference to the locutions of classical

Chinese. One might ask a native speaker to read aloud in the pronunciation of his own dialect a 'compound' from the lexicon of T'ang dynasty prose, or one might follow a dictionary interpretation of the structure of the classical phrase, but neither procedure, it seems, would reveal unerringly the solution to the question 'was this construction analogous to English *blackbird* or to *black bird*?' For this reason, the author's distinction between 'bound' and 'free' forms, which has proved so fruitful in the analysis of actual utterances, appears arbitrary and personal: 'The syllable represented by the graph [5376]¹ may, in this case, be translated "radical", but it cannot be used in isolation to represent a word "radical". This situation is described by saying that the syllable is *bound*, and not *free*. A bound syllable must occur in conjunction with other syllables, often with one particular other syllable, in order to be meaningful in a given sense.' I think that this analysis assumes that the conventional assignment of a given English word in translating the Chinese morpheme in a special environment constitutes a sufficient criterion of the freedom (or reverse) of that morpheme. It would appear more true to describe the situation thus:

(1) The 'word' [5376] corresponds quite well to English 'section' or 'department'.

(2) The 'word' [5376] functions freely (in literary Chinese) as an independent 'noun', as an independent 'verb', and as qualifier of 'verbs' and 'nouns'. Cf. such expressions as *pu chū* [5376, 1535] 'reside sectionally, or, in sections', and *pu shu* [5376, 5867] 'sectionalize and station'.

(3) The 'word' [5376], when preceded by one of a stated set of symbols used to separate the sections of a dictionary, is by convention translated by English 'radical', though the translation might just as well be 'section' or 'department'. In other words, the expression *tzu pu* [6939, 5376] 'child section [of the *Tz'u hai*]', sometimes Englished as "the child radical" is perfectly analogous to the expression *ping pu* [5282, 5376] 'arms section [of the old administration]', and even to *ch'i pu* [525, 5376] 'that section'. To say that [5376] is bound in the sense of 'radical' means only that 'radical' is a conventional (though by no means the best) translation of [5376] in some contexts.

An excellent case could be made however, for the propriety of the term 'bound' as applied to the elements of such expressions as *p'i-p'a* [5157, 4858], *chieh-yü* [788, 7602], *ch'ih-ch'u* [1026, 1402], and other binoms. All too commonly dictionaries, both in Chinese and in English, treat these elements as free and independent words. Here, however, the unit [5157] for instance is more than bound, it is meaningless, and not properly even a morpheme.

In telling of the twelve major divisions of the *Tz'u hai*, traditionally designated by the special names of the twelve zodiacal animals in a fixed order, Kennedy says, 'these signs are untranslatable' and 'Their use in modern Chinese is precisely parallel to our use of the letters of the alphabet'. The relevance of their use 'in modern Chinese' (the language of the man on the street?) to classical

¹ In this review, but not in the *ZH guide* itself, Chinese words are transcribed in the standard Wade-Giles romanization, followed in square brackets by the number of the graph with which it is written in Mathews' *Chinese-English dictionary*. Occasionally the bracketed number alone represents the word.

Chinese is doubtful. It is true that the series is memorized for the purpose of labelling things that come in regular sequence, especially hours, days, and years, but to call them untranslatable signs is incorrect. They are words in Chinese, and although their content (other than the mere connotations of 'usage as ordinal in such-and-such position') is weak, it is present. Doubtless every reader of a literary text is aware that *ch'en sui* [336, 5538] is a 'draconine year', even though [336] is not an everyday word for 'dragon'.

2. DATES. This invaluable section deals with far more than the several traditional chronological systems employed by the Chinese. The question of posthumous Imperial titles, for instance, is fairly dealt with. Kennedy, however, although he defines the word *tsung* [6896] as 'ancestor' and explains that it is prefixed by various laudatory though stereotyped epithets, does not go so far as to oppose the rather feeble-minded custom of leaving these titles untranslated, a custom which characterizes almost all publications about China. This doing-nothing tradition seems to be implicitly defended here on the specious grounds that the traits of character justifying the posthumous epithets are rarely discernible in the monarchs to whom they were applied. In short, the fact that a man was called 'the Venerable Ancestor' by his imperial descendants ought to be concealed from students of Chinese if his life-history shows him to have been anything but venerable. There are Chinese books which give ideal descriptions of a sovereign to whom a particular 'temple name' is appropriate. Kennedy seems to take these descriptions to be definitions of the epithets used in the titles. So it is made to appear that the epithets themselves are vague or virtually meaningless. Such is not the case. The Chinese tomes concerned with posthumous titles say in effect that the name 'Divine Ancestor' (*shen tsung* [5716, 6896]) applied to a deceased ruler implies that he was 'one who is unique and whom the people cannot name'. Granted that this description may be quite inappropriate to the monarch so styled, he was nonetheless clearly named 'Divine Ancestor', and we ought not to infer that [5716] is a word of imprecise reference. A translator is concerned with words and their meanings, not with the moot question of their correspondence to facts, although such matters may be treated in commentaries on the text.

The part of this chapter which deals with the common and sometimes complex method of dating events by reference to imperial reigns is very clear and precise, and ought to be studied carefully by any student of Chinese history.

3. PERSONALIA. This is a short chapter treating the basic biographical material found in standard Chinese reference works: personal names, expressions for birth and death, and the like. Especially valuable are two tables, one listing the common graphs employed as phonetic symbols for the transcription of foreign names, the other listing the most common Chinese surnames. Lack of familiarity with these graphs has too often made the interpretation of Chinese texts a nightmarish puzzle to elementary students. Here as elsewhere Kennedy displays his genius for recognizing exactly what the sinological neophyte ought to know. 'First things first' is a rule which has not until now had even lip-service from the authors of textbooks on literary Chinese.

4. BIOGRAPHIES. This is really an extension of the preceding chapter, and

deals with the linguistic formulae for the essential events in a man's life, especially his official career. Perhaps it is appropriate at this point to note that the *ZH guide* reproduces actual pages from the *Tz'u hai* from time to time, along with skilfully designed exercises that bring the student immediately to grips with fundamental problems, such as those designed to accompany the biographies reproduced here.

At the risk of appearing captious, I wish to comment on a few of the words defined in this chapter. *shih* [5776] 'lower official, squire, scholar, warrior': The practice of 'defining' a word by listing various kinds of persons (or situations) to whom it might be applied, though encouraged by tradition, seems to me of doubtful value. It also confuses the student, who naturally supposes that we can never really be sure just what the word means. A better practice would be to discover an English expression with approximately the same range of connotation or application as the Chinese, and use it as a definition. Thus: *shih* [5776] 'gentleman' (in the feudal age gentlemen were mainly knights; in later times the term was generally applied to members of the educated upper class; etc.) *teng* [6178]: Kennedy explains the usage of this difficult word very clearly, but does not suggest any useful translation, being confronted with the fact that 'A, B, C, [6178]' may mean 'A, B, C, and others' or 'A, B, C, these three'. An English expression of comparable ambiguity is 'the like of A, B, and C'. *wen* [7129] 'literature': Kennedy knows of course that 'literature' is only one aspect of [7129], but it would have been more consistent, perhaps, with his overall principle of 'honesty, always, even with beginners', to indicate that [7129] is a word of broad significance, meaning 'designed or patterned, especially pertaining to culture or the civil arts'.

5. PLACE NAMES. Despite the disclaimer in the Foreword of any exclusive or overemphatic preoccupation with recent times in a textbook for general sinologists, the author has in this chapter (and later in the chapter on Positions) slanted his illustrations and explanations somewhat in the direction of Ch'ing dynasty and Republican institutions, as if they were either typically Chinese (whatever that may be) or somehow more important for the student of Chinese civilization than those of earlier times. For instance, as examples of Chinese words applied to areas of the provincial administration, he has chosen *sheng* [5744], *fu* [1928], *hsien* [2700] and *ch'eng* [380], but omitted mention of *chou* [1289] and *chün* [1718], both of which were of fundamental importance through most of the history of the Empire. The earnest student of the Chinese language and the history of Chinese ideas might also benefit from further information about these administrative expressions. Kennedy has given stereotyped English equivalents: [5744] 'province', [1928] 'prefecture', [2700] 'district', [380] 'city'. A book aimed at developing 'philological' habits of inquiry in the student might well have added:

[5744] 'originally a Han dynasty substitute for the then tabooed word *chin* [1077] "sacred, forbidden", in the sense of "pertaining to the sacred palace of the Son of Heaven, and to the offices attached to it or deriving their authority from it"'

[1928] 'storehouse, repository, library; applied to the offices of certain officials'

[2700] 'suspended from; dependent on; because attached to larger administrative districts'

[380] 'city-wall; hence, a walled city' ,

6. JAPANESE NAMES. Here Kennedy, with his usual clarity and good humor, has outlined the complexities which have resulted from the Japanese use of Chinese graphs to represent not only Chinese loanwords (or morphemes), but also groups of native Japanese words thought to be more or less synonymous with the Chinese words for which the graphs originally stood. The chapter includes also a useful table of graphs commonly used to represent parts of Japanese proper names together with their readings, which should be of great value to the sinological student who is relatively unsophisticated in things Japanese.

The two sources suggested as helps in deciphering Japanese names, Rose-Innes and Papinot, are unfortunately not too reliable. The notion that Chinese graphs can be looked up in Rose-Innes 'in the usual manner' is altogether misleading because of that authority's singular arrangement of the radicals. It might better have been suggested that the student learn the kana syllabary (not too difficult a task), after which he can find the proper readings for a Chinese graph in Ueda's *Daijiten*.

7. WARFARE. In this chapter the author takes advantage of the fact that 'verbs form the chief part of the vocabulary of warfare' to analyse in some detail the functions of the 'verb' in literary Chinese, especially in relation to its 'object'. This analysis is of basic importance to the understanding of Chinese syntax, and Kennedy has given an admirably succinct and intelligible account of the situation. In addition, he has made a significant contribution to the 'sociology' of language by indicating the moral directivity of the lexicon of battle. Thus: *cheng* [352] 'invade, attack (righteous action)'; *fan* [1779] 'invade (enemy action)'. This kind of information is essential to the student, and rarely provided by the standard dictionaries. Your reviewer believes, however, that greater precision might have been used in the definitions, thus: *cheng* [352] 'go on a CORRECTIVE (cf. *cheng* [351] "correct") march or expedition [against an enemy]'; *fan* [1779] 'violate (a law, sacred ground, etc.) [used of an enemy invasion]'.

8. LEXICAL FORMULAE. This is another invaluable section, treating the various standard devices whereby Chinese dictionaries define words. Facility at understanding such definitions is normally picked up empirically and painfully by the students of Chinese, since the writers of Chinese 'grammars' and the teachers of the literary language in the past have been too much concerned with imparting uncritically the traditional interpretations of the classics to bother about the analysis of the apparatus of word and phrase definition. In general the reverse is true of this book: Kennedy supplies first what is basic—an almost unheard-of procedure.

I question, however, the complete validity of his description of the important word *yeh* [7312], which he describes as a 'pause-marker'. It is indeed a pause-marker, but it is also a word in Chinese, although from Kennedy's account of it one could easily get the impression that it is analogous to such markers as '?' and ':', a non-lexical sign suggesting some kind of sentence intonation or open juncture. To say in English what the word *yeh* [7312] means is of course not easy.

It looks to your reviewer like a kind of demonstrative, but at any rate an approach to its meaning might be more satisfactorily achieved by comparing the 'definition' sentences given by Kennedy with sentences in which *yeh* is used to point up contrasted subjects, and sentences (containing a verb) in which it emphasizes the finality of a judgment.

9. DERIVATIONS. Here is further grammatical information, having to do with the most important 'particles' or 'adjuncts' of literary Chinese: *so* [5465], *che* [263], *erh* [1756], and *ch'i* [525]. Almost everything that needs to be said about these words is stated in a few short paragraphs, an admirable feat of lucid condensation. Also contained in this chapter, somewhat adventitiously, is a table of customary short names for Chinese provinces, together with their longer equivalents and pertinent translation exercises. This again is material which ought to be learned during the first year of Chinese, but all too seldom is.

10. POSITIONS. Beginning with a brief account of the class of words which he styles 'localizers' (Chinese 'nouns' translatable by 'upper', 'under', 'in front of', etc.), Kennedy ranges through such problems as meaning and form (or position) as criteria for word-classes, the uses of the important coverb *i* [2932], and titles in the Chinese bureaucracy of the Ming and Ch'ing dynasties. Everything said here is cogent and valuable, and your reviewer does not object to the compressing of such apparently diverse subject matter under one heading. A few special points, however, need to be made with reference to these topics.

It might have been desirable, after a clear-cut discussion of the applicability of such terms as 'adjective' to any set of words in classical Chinese, to outline what classes of words can in fact be determined in that language by virtue of their position. I have found, for instance, at least fifteen possible positions in the 'ideal' or 'completest possible' Chinese sentence, and words might be described adequately by taking note of the positions they may occupy. The positions can be designated simply by numbers, or by arbitrary grammatical expressions. Thus 'position I' means the same as 'pre-subject expressions'. Some words are limited to this position (as *erh* [1756] and *tan* [6038]), and might be styled 'exclusive pre-subject expressions'; some are found also in other positions (as *chin* [1053] and *tse* [6746]), and might be styled 'occasional pre-subject expressions'. A sample description of a word for a *Literary Chinese dictionary* of the future:

[7615] *yü*² (1) 'eh?' (interrogative final particle). Position XV.

*yü*³ (2) 'with; and' (formative for correlative noun phrases). Positions II, V, XI, XIII (i.e. all positions which define the category 'noun'), always preceded and followed by simple 'nouns'.

(3) 'together with' (secondary verb). Position IV, with or without noun 'object'.

(4) 'give' (verb). Position X.

*yü*⁴ (5) 'participate in' (verb). Position X.

Such a system would present adequate criteria for defining the 'parts of speech' in terms of their regular syntactical relationships to other words in the sentence,

and would avoid the difficulties accompanying a fixed classification of lexical items as 'verbs', 'nouns', etc.

In reference to the official titles discussed in this chapter, I have already in another context suggested that Kennedy has laid undue emphasis on recent institutions and the vocabulary which describes them, in spite of his prefatory statement that his book is designed for persons who lack a bias in favor of modern political history. Furthermore, in his selection of quoted examples, he has given tacit support to the all-too-widespread dogma which makes Chinese titles (and other expressions) 'intelligible' to the Western student by translating them with familiar titles chosen from the political systems of Western Europe and the United States. In my opinion, this is to seriously distort the Chinese text, loading the translation with connotations completely foreign to Chinese thought and custom, and above all to the Chinese language. Thus Mayers' versions, reproduced in this book without critical comment, include: *shang shu* [5670, 5857], translated 'President of a Board', and *shih lang* [5778, 3820], translated 'Vice-President of a Board'. Such versions neglect the meaning of the Chinese altogether, taking a familiar English title on the grounds that it is or was held by the chief official of some branch of government. This is as if one were to translate Arabic *amīr al-mu'minīn* by 'president', or 'dictator', or 'pope', depending on one's belief about the role of the holder of that title at a given period of Islamic history, instead of rendering the LANGUAGE of the text: 'Commander of the Believers'. For the two examples cited above I would suggest: *shang shu* [5670, 5857] 'Provost Scriptural', and *shih lang* [5778, 3820] 'Esquire Officiant'. The FORM of these translations is analogous to *Princess Royal*, *Governor General*, *Heir Apparent*, and so forth. As to meaning, they render, with some semantic fidelity, the literary Chinese words *shang* 'prefer; prepositorial; provost (one placed before)'; *shu* 'writing; scriptures, etc.'; *shih* 'officiate; official, etc.'; *lang* 'young gentleman; esquire'. These meanings are and were alive in the written language. In an age when the unquestioned tendency to write history in terms conformable to local prejudices is rampant, the old spirit of philological precision needs to be revived and encouraged. A serious translator has other duties than to provide easily assimilated substitutes for 'difficult' and strange expressions.

11. CITATIONS. This is a short chapter about dictionary citations which illustrate the uses of words and phrases by referring to the ancient classics. Understanding these references is a *sine qua non* of sophistication in Chinese lexicography.

12. READINGS. This chapter deals with phonology and orthography, fields in which Kennedy is an acknowledged master. Of particular interest and of the greatest merit here is Kennedy's simplified spelling of Ancient Chinese (used throughout the volume). It is notorious that Karlgren's romanization of the language of the T'ang dynasty makes use of a bewildering number of diacritics and special symbols. Kennedy states, to justify his modification of this spelling: 'The purpose of an orthography is to transmit language for the eye, and it fails to do this unless its forms are easily recognizable, easily remembered, and easily reproduced by the familiar instruments of writing and printing.' Nothing could

be more true, and the 'Kennedy system' or something like it has been long overdue. Unfortunately, although the *ZH guide* contains an index of words with readings given both in 'Karlgren' and in 'Kennedy', this does not provide instances of all of the morphemes of Ancient Chinese; and no systematic explanation of how Kennedy's (phonemic?) spelling is derived from Karlgren's phonetic spelling is to be found. Such a presentation of the new orthography as a part of the *ZH guide* would greatly enhance its value.

In conclusion, this book is highly recommended for students of literary Chinese on whatever level.

Dictionnaire ngbandi (Ubangi—Congo belge). By P. BENJAMIN LEKENS. (Annales du Musée Royal du Congo Belge; Sciences de l'homme: Linguistique, Vol. 1.) Pp. xii, 348. Anvers: Éditions de Sikkel, 1952.

Dictionnaire français-lomongo (Lonkundo). By H. HULSTAERT, M.S.C. (Annales du Musée Royal du Congo Belge; Sciences de l'homme: Linguistique, Vol. 2.) Pp. xxxii, 466. Anvers: Éditions de Sikkel, 1952.

Reviewed by JOSEPH H. GREENBERG, *Columbia University*

These two dictionaries furnish welcome evidence of the high quality of contemporary Belgian linguistic work in the Congo. Both of them definitely supersede previous lexicographic work on the languages concerned, and represent major contributions to our knowledge of African languages.

The language treated by Lekens, Ngbandi, belongs to the eastern branch of the Niger-Congo family, along with Zande, Gbaya, Banda, and a number of other forms of speech; it is spoken in the central basin of the important Ubangi tributary of the Congo River in the Northern Belgian Congo. A widely distributed dialect, Sango, extends into neighboring French Equatorial Africa. It is employed as a trade language throughout the area; in the words of G. H. Brachiel, *Vocabulaire Sangho, Haute Oubangui* 3 (Paris, 1909), it is 'une espèce de Volapuk'. The writer of the present dictionary has had extended experience with the language and is the author of the standard grammatical treatise, *Spraakkunst der Ngbanditaal* (Brugge, 1923).

With a few minor modifications, the orthography in Lekens' dictionary is phonemic. Almost all the problems refer to the labiovelars *kp* and *gb*. These are evidently nondistinctively labialized before front vowels, and followed by a nondistinctive [m] before nasalized vowels. Thus Lekens' *kpwe* 'men' is phonemically /kpe/, and *kpmĩ* 'fish' is /kpĩ/. Phonetic details are rather sparingly given in the author's introduction, but fortunately we have the excellent sketch of Burssens, *De Klinkerphonemen in het Gbandi* 257-70 (Kongo Overzee, 1936), to supplement the information given here. The language described by Burssens appears to be identical with that of Lekens' dictionary.

This is one of the relatively few dictionaries of African languages that has the tones marked throughout. There are a number of proofs of Lekens' accuracy in this respect: the complete agreement with Burssens's tonal notations, the complete agreement with Nida's markings in his *Morphology* (if we make allowance for the obvious dialect differences), and the substantial agreement on a com-

parative basis with the forms of related Niger-Congo languages: of 17 cognates with Proto-Bantu noted in passing, 13 agreed in tone.

The second of the works to be considered deals with Lomongo, often called Lonkundu, a member of the extensive Bantu subfamily of the Niger-Congo family. Lomongo is spoken by perhaps as many as 2,000,000 people in numerous dialect variations, well to the south of Ngbandi in the central basin of the Congo River. Hulstaert's dictionary is based on the speech of the northwestern part of the Lomongo area, specifically that of Bonkoso-Bongili. One weakness of the present work is the quotation of forms of other dialects, with the notation D but without further indication of regional provenience. There are several dictionaries of Lomongo in existence, notably the extensive work of E. A. Ruskin, *Dictionary of the Lomongo language* (London, 1928). Hulstaert's work is superior in that it is completely phonemic, marking both tone and the distinctions between *e* and *ɛ*, *o* and *ɔ*, not found in Ruskin's work. In addition, there is a brief but valuable introduction giving the tones of bound grammatical elements. Unfortunately this is a Lomongo-French dictionary only; but a Flemish-Lomongo lexicon is announced as in press.

As the first works in the new linguistic series published by the Institut Royal Colonial Belge, the present dictionaries mark a most auspicious beginning to what promises to be a valuable series of contributions to African linguistics.

Malgache et maanjan: Une comparaison linguistique. By OTTO CHR. DAHL. (Avhandlingar utgitt av Egede-Instituttet, Vol. 3.) Pp. 408. Oslo: Egede-Instituttet (Arne Gimnes Forlag), 1951.

Reviewed by ISIDORE DYEN, *Yale University*

The purpose of this book is to determine the 'kind and degree' of relationship between the Malagasy language of Madagascar and the Maanyan language of southeastern Borneo (27); both languages are Malayo-Polynesian. The position of Malagasy among the Malayo-Polynesian languages is a curious one, and never fails to arouse interest. Its obvious membership in the Malayo-Polynesian family and its geographical isolation lead immediately to two questions: How did the speakers reach Madagascar? and Can we identify the language of these settlers more closely than simply as Malayo-Polynesian?

A major observation which conditions all linguistic discussion of Malagasy is its linguistic unity. The dialects can be divided into two large groups: the western (Sakaláva, Vézo, Mahafály, Tanalána, and Tándroy on the west coast) and the eastern (Tanósy, Taimanambóndro, Taisáka, Saháfatra, Taifásy, Taimóro, Tambahóaka, and Betsimisáraka on the east coast; Sihánaka, Bezanozáno, and Tanála in the eastern watershed; Bára, Betsiléó, Mérina, and Tsimihéty on the high plateaus), with some transitional dialects (Tankárana in the north and to some extent Tanósy in the south). The dialects differ somewhat in phonetic evolution and particularly in vocabulary, whereas their grammar is remarkably uniform. 'L'unité linguistique de cette grande île, une des plus grandes du monde, est étonnante en comparaison avec l'état linguistique des îles semblables de l'archipel Indonésien. Cela fait supposer que Madagascar a été

peuplée par une population parlant malgache depuis une époque relativement récente' (5-6).

1. Dahl gives a brief résumé of the history of the Malagasy and of the study of their language (5-23). This is followed by a similar treatment of Maanyan (23-6), material for which is limited to two vocabularies, one folkloristic and a few religious texts, and a brief grammatical treatment. Dahl finds that in general the religious texts are trustworthy representatives of the language, since they are much like the fables and tales.

The main body of the book is devoted to a linguistic comparison, which reaches the following conclusion (355): 'Le maanjan et le malgache sont des langues très étroitement apparentées. Parmi les langues indonésiennes connues actuellement aucune n'a autant de ressemblances avec le malgache que le maanjan.' We will return to the linguistic comparison and its significance below; here we will turn to Dahl's accumulation of historical evidence.

Dahl (356 f.) excludes the possibility that the Malagasy arrived in Madagascar directly from the East Indies in boats that were blown off course and shipwrecked. There is linguistic evidence for supposing an autochthonous population, and the language of shipwrecked mariners would not have persisted. Furthermore it is unlikely that an unprepared crew could have survived such a voyage.

The historical evidence bearing on the source of the Malagasy, derived mainly from the research of G. Ferrand, is based on Arab sources. Two such sources, of the 11th and 12th centuries, speak clearly of navigation between Sofala (once an important port of Mozambique) and China and the Indies. The navigators followed the coasts of India, Arabia, and Africa, for they stopped at Kathiawar and Aden.

The author places great stress on two geographical names: *Qumr* and *Wāqwāq*. He believes Ferrand has proved that the island *Qumr* near the African coast is Madagascar (358). A people called *Qumr* are reported by Ibn-al-Muğawir, an Arab author of the 13th century—according to a local tradition—to have taken possession of Aden after the decline of the empire of the Pharaohs. Dahl thinks that this more probably occurred from the 3rd to the 5th century with the decline of the Roman Empire, because people would look for new markets after the traditional commerce with Europe was reduced (364).

There are however two *Wāqwāq*'s; according to one Arabic author, one was near China and the other in 'yaman'. The author does not attempt to explain this word (which is usually identified with Yemen), since another Arabic author places the western *Wāqwāq* near Sofala. 'Il est donc certain qu'au 10e siècle, on connaissait deux *Wāqwāq*, l'un à l'est et l'autre à l'ouest de l'Océan Indien' (358). Ferrand identifies one *Wāqwāq* with Madagascar by equating the name with Mlg. *vahodka* 'people, subjects (in relation to a sovereign)'. Ferrand saw here a word of Bantu origin, whereas Dahl believes that it is an Indonesian word and that 'le nom indonésien des immigrants à Madagascar fut *wakúak' (359).

Dahl identifies the eastern *Wāqwāq* with Borneo, against those who have identified it with Japan (de Goeje) or with Sumatra (Ferrand). The statements of Arabic authors place this *Wāqwāq* in the China Sea near the islands of Zābag, identified as Sumatra. 'Pour les marins qui connaissaient l'Extrême Orient,

Wāqwāq était l'archipel qui se trouvait à l'est de leur route de Sumatra à la Chine ... en premier lieu les îles Philippines et Bornéo (363). Since Borneo was the nearest of the islands called Wāqwāq to the route followed by the Arabs, 'il est donc possible qu'il y a ait eu [sic], à Bornéo, un peuple ou un état célèbre dont le nom a été identifié aussi bien avec le Wāqwāq oriental qu'avec le Wāqwāq insulaire africain' (ibid.).

In 945/946 A.D. an Arabic navigator observed a great Wāqwāq fleet pillaging in the Mozambique Channel, 'probablement à Madagascar, en cherchant des marchandises convenant à leur pays'. Dahl concludes that it has been proved that 'il y a eu relations maritimes entre l'Indonésie et Madagascar, et que ces voyages furent entrepris par des gens de Qumr et de Wāqwāq. Ibn-al-Muğawir nous informe aussi que leurs navires étaient à balanciers, c'est-à-dire des embarcations indonésiennes' (364).

Dahl assumes that the exploits of the Qumr of the Wāqwāq fleet in their respective times were associated with the control of trade routes. Relations between Sofala and Sumatra probably still continued in the 12th century, but the power of the Qumr must have ended by the 13th century and the route closed. A diminution of Indonesian commerce in the 12th century is probable, because this was a period of Mohammedan expansion in Madagascar. The author leaves open the question whether such relations did not continue down to a more recent period.

The date of the first migration is set by Dahl at approximately 400 A.D. Old Javanese *jěnggī* 'a kind of slave' (identified by H. Kern with Malay *djanggi*, Toba-Batak *djonggi* 'African Negro') is connected with Arabic *zanġ* 'East Africa', *zanġī* 'East Africans', and with the Chinese *seng-k'i* or *seng-tche*, the name of some slaves offered at the court of China in 724 and 813. This hypothesis is confirmed by Chinese texts that speak of a country named *K'ouen-louen ts'eng-k'i* 'in the sea of the southwest'. This country is certainly the East African coast, whereas there is another *K'ouen-louen* in Indonesia. This double reference is therefore a parallel to the two Arabic Wāqwāq (366-7).

Before 724 A.D. there is no historical evidence bearing on the date of the first migration. Words of Sanskrit origin are all, with one exception, found also in the languages of the East Indies; sometimes the same peculiar phonetic and semantic irregularities reappear in relation to the Sanskrit words. The number of such words in Malagasy, however, is comparatively small. The only possible explanation of this phenomenon is that the Indonesian immigrants left their country of origin at the beginning of the Hindu period. Since that country was Borneo, we have a means of dating the beginning of this influence and thereby the emigration to Madagascar (367).

The oldest inscriptions in Indonesia, found at Muara Kaman in eastern Borneo, are placed at 400 A.D. They were made by Brahmans in honor of king Mūlavarman, son of Aṣṭavarman and grandson of Kuṇḍunga; the latter seems to have had an Indonesian name. The inscriptions thus date from the beginning of Brahman influence in Borneo.

Beside the religious influence of the Sanskrit used by Brahmans, there was certainly a lay influence of the Prakrit used in commerce. Note the commercial

character of the Indic vocabulary in Malagasy, and the commercial meaning of the word *trosa* 'debt' (Skt. *doṣa* 'sin'). In this way Dahl arrives at dating the migration which bore the words of Indic origin to Madagascar toward 400 A.D. (368).

2. The linguistic problem which Dahl undertakes is of a type which has not usually been assigned a formal term in linguistics. I use the term 'subgrouping' for such procedures as are employed to prove the closer relationship (or sub-relationship) of two or more languages of a family. A few words discussing such procedures are in point in connection with an evaluation of Dahl's work.

Three subgrouping procedures are now available to the comparatist: (1) judgment by inspection; (2) discovery of exclusively shared (non-accidental) innovations; (3) lexicostatistic dating.

A procedure for subgrouping is set forth by Brugmann as follows:¹ 'it is not a single or a few linguistic phenomena appearing in two or several areas at the same time which furnish a proof of closer community, but only a large mass of agreements in sound, flexional, syntactic, and lexical innovations, the large mass of which excludes the thought of accident.' The term 'innovation' as used here could be satisfactorily defined as any linguistic difference between a proto-language and a daughter language. Brugmann's statement excludes 'retentions' as criteria of closer relationship. This exclusion appears obvious and necessary to the purpose; for if we were to consider retentions, we should be using the criteria of proto-relationship also as evidence of subrelationship (or closer relationship). It will then follow that the proto-language must be reconstructed exactly, or at least clear criteria must be available for determining whether a given shared linguistic feature is a retention or an innovation.

In contradistinction to Brugmann's usage, closer relationship can be used as a broad term to include both membership in the same dialect area of the proto-language as well as membership in the same subgroup, where a subgroup implies a private proto-language which was one of the languages resulting from the differentiation of the proto-language of the whole family. In this sense a closer relationship is indicated by any exclusively shared innovation.

A shared innovation is one which cannot be due to chance (i.e. to independent linguistic change) or to separate borrowing. We speak of exclusively shared innovations because a conclusion drawn from a shared innovation applies equally to all languages that share the innovation.

The need to consider more than one exclusively shared innovation in establishing a subgroup arises when we attempt to demonstrate that a language boundary appeared in the dissolution of the proto-language of the family, a boundary that included what today appear as different languages. It is of course conceivable that a single exclusively shared innovation is the sole remnant of such a boundary. But the demonstration that such a boundary actually existed depends generally on the appearance of a set of exclusively shared innovations; the mass of such innovations, however, need only be greater than that which could be due to the dialectal spread of innovations (whose borders now coincide as a result of later

¹ Zur Frage nach den Verwandtschaftsverhältnissen der indogermanischen Sprachen, *Internationale Zeitschrift für allgemeine Sprachwissenschaft* 1.253 (1884).

developments) within the proto-language. Thus the mass and the identity of range of sets of exclusively shared innovations in related languages determine the probability of the former existence of a language boundary.

Distinctions between some innovations and some retentions can be observed even under conditions where not all retentions can be distinguished from all innovations. Such distinctions may be sufficient for the construction of a subgroup. If any two or more related languages share a feature, the question arises whether this is a retention or an innovation. If we apply a general rule that such features are taken to be retentions unless there is evidence to the contrary, then a corresponding proto-feature is reconstructed. It follows that (borrowings being excluded) an innovation occurring in two or more languages can be detected only if, as a proto-feature, it contradicts a proto-feature which for some reason appears to be more ancient. As soon as such a discovery is made and it is clear that independent linguistic change (accident) is not a satisfactory explanation, we posit a closer relationship (as defined above) of the languages that exhibit the result of the contradictory less ancient proto-feature. Such features are exemplified by some of those usually taken to indicate isoglosses within Proto-Indo-European (cf. Bloomfield, *Language* 316). If, however, contradictory less ancient features occur over the same range of languages to an extent sufficient to rule out explanation as dialectal developments in the proto-language, we posit a subgroup. The discovery of such contradictions is easiest where they are systematically related, as in the morphology. The verbal system of the Germanic languages is different from that of Vedic Sanskrit; the latter, on the other hand, can be viewed as suggesting a system from which both the Germanic and the Greek verbal systems arose. We conclude that the Germanic languages constitute a subgroup. Just as a band of isoglosses constitutes a language boundary, so a set of exclusively shared innovations determines the former existence of a language boundary.

If Brugmann's statement is now re-examined, it is evident that for him the main problem is to rule out chance as a factor in the determination of subgroups, and so it is if indeed innovations are viewed as isolated data. For any given instance of a shared innovation there is always the chance that independent linguistic change is involved. Brugmann's conclusion is that we can construct a subgroup only if the innovation of all types is great enough. But the double requirement that only innovations be considered and that the mass of innovations (taken in isolation) be great enough, may very well yield no more than can be obtained by simple inspection of the languages. In the view presented above, chance (or independent linguistic change) is taken as calculated in each item and controlled for sets of innovations by the demand that their distribution among languages be identical.

It appears that in the field of Indo-European the procedure of subgrouping was carried out without reference to Brugmann's dictum: it was done with direct reference not to a distinction between retentions and innovations, but only to the variety and degree of detail in the resemblances. Thus the Germanic languages resemble each other in many features and differ in those features from other languages, regardless of the question whether those features are retentions

or innovations. The multiplicity of shared differences from other languages seems to be so great that the operation of subgrouping those of the Germanic group could be performed immediately. In the field of Malayo-Polynesian, I believe that we can similarly subgroup Tagalog, Bisayan, and Bikol. This procedure can be called 'subgrouping by inspection'. It depends for its validity on the probability that some of the exclusively shared differences are not due to independent linguistic change, separate borrowing, or retention.

Brugmann's criterion appears thus to be a mixture of what has been called here judgment by inspection and the discovery of exclusively shared innovations. To the extent that innovations cannot be demonstrated, the comparatist who attempts to follow out the implications of his criterion is forced back upon judgment by inspection.

Dahl's procedure illustrates the point, though to be sure he makes no direct reference to Brugmann's criterion. He collects the shared phenomena of Maanyan and Malagasy, indicating those which are shared exclusively by these two languages and those which they share with other languages of Borneo, and reaches the conclusion that these phenomena are sufficient to indicate the subrelationship of the two languages. The differences between the languages are then quite properly (on the basis of the hypothesis reached) assigned to the post-separation period of each language. Since, however, systematic conclusions are available only regarding PMP phonemes, it is only in dealing with the phonemic comparisons that Dahl can distinguish between innovations and retentions. His procedure with respect to the comparison of other linguistic features resembles most closely that of judgment by inspection. The first conclusion that I reached was that Dahl's material is not convincing; though his hypothesis is interesting and the material collected is sufficient to warrant further research, he has not proved the closer relationship of the two languages.

Dahl proceeds systematically. He divides his comparative procedure into three parts: (1) *comparaison phonétique* (28-119); (2) *grammaire* (120-298); (3) *vocabulaire* (299-355). It is clear from his discussion that he considers the first part sufficient and the third part confirmatory, whereas he finds no good evidence in the second part. We shall deal with the second part first.

3. Dahl concludes the section on grammar (297 f.) in this way: 'Vu les grandes convergences dans les différentes catégories de mots pleins, il est étonnant de constater combien différent les mots accessoires. Mais il ne faut néanmoins pas en tirer la conclusion que ces catégories aient évolué entièrement après la séparation des deux langues. Nous avons maintes fois constaté que les idées qui ont servi de base pour l'évolution, sont les mêmes dans les deux langues, et également que des mots d'étymologies différentes ont un emploi identique. Il est donc évident que nous avons beaucoup de cas où de nouveaux mots se sont substitués à d'autres, ayant eu la même fonction auparavant.' It is fair to say that Dahl finds that this part of his investigation does not strongly support his case. He adds, however, that (398) 'on trouve partout de petits détails identiques, communs aux deux langues seulement, détails qui ne peuvent être fortuits.' Such details are insufficient without a demonstration that they are shared innovations, i.e. that they do not reflect features of the parent language and are not due to

coincidence; to demonstrate that they are innovations would necessitate a proof that these shared differences contradict a feature which can be plausibly assigned to the proto-language. Dahl does not enter into such a demonstration, and it is for this reason that his procedure resembles that of subgrouping by inspection. To the question whether the totality of shared grammatical differences is by itself convincing, Dahl's answer appears to be no. We might ask whether the totality of shared grammatical differences adds much weight to the argument for a closer relationship; I believe that a fair reading of Dahl's chapter and of his conclusions will also yield the answer no.

Some of the shared vocabulary items are very striking, e.g. Mny. *kariwe*, Mlg. *hariva* 'afternoon, evening' (cognates in other languages are so far unknown). Any argument from unclassified vocabulary must meet one of two tests: (1) at least one case must involve a shared innovation; (2) the number of items must be large enough to exclude the possibility that they are all shared retentions from the earlier proto-language (and of course the factor of borrowing must also be somehow excluded). Actually we should need only a single item that meets the first test. But it is a peculiarity of the comparative method in its present rudimentary state, that whenever we reconstruct an etymon for material from languages so far apart geographically as Maanyan and Malagasy, the conclusion we reach is just as satisfactory if we assign the reconstruction to Proto-Malayo-Polynesian as it would be if we assigned it to a hypothetical proto-language peculiar to Maanyan and Malagasy. There is no way of demonstrating that a shared vocabulary item is a shared innovation unless (1) it contains a demonstrable morphological innovation, or (2) it exemplifies a phonetic change which is not likely to have recurred independently.

Dahl concludes his section on vocabulary as follows (355):

Nous venons de rapprocher environ 500 radicaux de nos deux langues. Il y a des rapprochements qui ne sont que probables, mais la plupart me semblent certains. Parmi ces mots la plus grande partie appartient au vocabulaire commun aux langues indonésiennes, et prouve seulement que les deux langues sont membres de ce groupe. Mais il y en a qui présente les mêmes irrégularités phonétiques par rapport à l'indonésien commun, et les sens évolués sont remarquablement convergents. De plus, il y a un assez grand nombre de mots que je n'ai observés que dans nos deux langues et d'autres qu'elles ont en commun avec les langues bornéennes seulement.

De la littérature maanjan, ainsi que des vocabulaires de den Hamer et de Ray, j'ai recueilli un peu plus de 2000 radicaux maanjan. C'est donc presque le quart de ce vocabulaire que j'ai pu identifier avec des mots malgaches.

None of the comparisons cited seems to be a positive instance of a shared innovation. Under judgment by inspection, our interest is limited to the number and type of the exclusively shared words. I have made the following hasty calculation. Of about 481 comparisons, about 326 seem to have no effect on the problem; they are those in which (1) the cognates are phonetically regular and occur in other than Borneo languages, with meanings not specially limited to Borneo, or (2) the comparison is expressly labeled by the author as doubtful. Of the remaining 155, about 25 are known to occur in other Borneo languages but are not known to occur outside Borneo, and 6 are known to occur outside Borneo

but with the given meaning occur only in Malagasy and Borneo languages. Seven show a feature which is peculiar to Malagasy and Maanyan. In about 110 comparisons, only Malagasy and Maanyan cognates are involved, and of these I find about 90 attractive. To be added are about 8 comparisons in which a word is found outside Borneo, but is found only in Malagasy and Maanyan with a particular meaning. The remaining comparisons I found difficult to fit under one rubric or another.

A great deal then depends on what attitude is taken toward the 110 + 8 words. The number is certainly large, but we may well ask whether the number will not be reduced as more languages are brought into comparison. Even for the putative remainder, of course, we should still be unsure how many were retentions and not innovations.

As for the type of words which enter into the comparison, we shall see that there is a striking number from the area which we can call 'basic vocabulary'. Dahl does not make much of a point of this. He classifies the compared vocabulary under the headings: *Le corps humain, Fonctions du corps, Famille, Maison et ménage, Terre, Eau, Animaux, Végétaux, Ciel et temps, Civilisation matérielle, Fonctions et réactions spirituelles, Qualités et relations diverses, and Mots traités dans la partie grammaticale*. His purpose in doing so is (299) 'pour montrer que les mots convergents appartiennent à tous les domaines de la vie humaine'.

4. Here is Dahl's estimate of what he found in his 'comparaison phonétique' (95 f.):

... l'évolution phonétique de ces deux langues est remarquablement semblable. Les complexes de consonnes au milieu du radical ont été simplifiés, et les occlusives sonores ont été assourdis en position final. On retrouve cependant ces phénomènes dans un assez grand nombre de langues très éloignées les unes des autres, et dont les autres caractéristiques sont assez différentes. Seuls, ces phénomènes ne sont donc pas décisifs.

Mais il y en a d'autres plus importants. L'*a* final est devenu *e* dans les langues du sud-est de Bornéo. C'est une deuxième convergence entre le malgache d'une part et le maanjan et ses voisins d'autre part.

Toutefois, le fait le plus significatif est l'évolution des phonèmes *d, ɖ, g'*, *l* and *ɣ*. Ces phonèmes ont donné des résultats très différents dans les langues indonésiennes. L'un et l'autre de ces phonèmes a souvent donné des résultats divergents même dans les langues qui paraissent étroitement apparentées. En maanjan et en malgache tous ces phonèmes ont les mêmes représentants. Certains d'entre eux ont même donné deux résultats qui, cependant, sont convergents dans les deux langues. Il en est de même avec le phonème *t'* dont l'évolution a abouti à *s* et à *h/zéro*. Remarquable est aussi le fait que *li* est devenu *di*, et *ti* est devenu *si/tsi* en maanjan et dans les dialectes orientaux du malgache.

En somme, l'évolution de tous les phonèmes des deux langues est la même jusqu'à un certain point. Depuis, c'est surtout le malgache qui a poussé son évolution plus loin.

Une telle convergence dans l'évolution de deux langues n'est pas possible si elles n'ont pas eu une partie de leur évolution en commun. Autrement dit, ces convergences nous forcent à supposer que le maanjan et le malgache sont des branches d'une même souche, ou plutôt des rameaux de la même branche.

Je ne connais aucune autre langue indonésienne qui ait la même évolution de tous ces phonèmes que nos deux langues. Autant que nous le sachions il n'existe donc pas d'autre langue plus apparentée à nos deux langues qu'elles ne le sont entre elles.

Même le lowangan, que les missionnaires de Bornéo caractérisent comme plus étroitement apparentée au maanjan que les autres langues voisines, a une autre évolution de *d* et de *ɖ*, généralement aussi de *ɓ*, et souvent de *ɣ* ... Cette langue a donc du séparer du maanjan à une époque plus éloignée que le malgache.

It is difficult to decide how Dahl weights the changes **li* > Mny. and eastern Malagasy *di*, and **ti* > Mny. *si*, eastern Malagasy *tsi*. He clearly indicates that he recognizes that these changes were completed independently. He says (112), 'Après la séparation des deux langues, la phonétique du maanjan a très peu changé. Les tendances de changement de ... *ti* et de *li* ont abouti à ... *si* et à *di*.' In that case these changes play no part in the argument for closer relationship. If closer relationship should be proved, these changes can be cited as instances of convergent 'drift', but such instances cannot be cited as evidence of closer relationship.

We may also discount with Dahl the changes involving simplification of internal consonant clusters and unvoicing of voiced finals. As for the remaining phenomena which Dahl finds important, we must distinguish two types: (a) those in which the phonetic results of a proto-phoneme are similar in the two languages but different from what is found elsewhere, and (b) those in which Dahl claims that a proto-phoneme has given different results in the two languages and these results are phonetically similar.

Under type (a) are to be classed the following. (1) Some **d* and **ɖ* and all **g'* > Mlg. *r*, Mny. *r* in positions other than final, where all **d*, **ɖ*, and **g'* > Mlg. (if not lost) *tr(a)*,² Mny. *t*; (2) **l* > Mlg. and Mny. *r* in all positions except final, where it > Mlg. (if not lost) *tr*, Mny. *r*; (3) Some **ɣ* > in non-final position > Mlg. and Mny. zero (whereas some **ɣ* in non-final position > Mlg. *r*, Mny. *r*; I omit the complicated developments in final position). The falling together of some **d* and **ɖ* and all **g'* in all positions is an interesting phenomenon. Furthermore it is interesting that the result falls together with that of **l* and with some **ɣ* if the results in final position be ignored. The importance of this observation would perhaps be greater if it were not also observable that **d*, **ɖ*, **g'* > Malay *d* in all positions except final, where **d*, **g'* > Ml. *t* and **ɖ* > Ml. *r*, and that **l*, **ɣ* > Ml. *r* in all positions. The difference in result in Malagasy and Maanyan from the result in Malay in non-final positions could be summed up as a change **d*, **ɖ*, **g'* > **d* > Ml. *d*, Mlg. and Mny. *r*. The strength of Dahl's argument thus appears to lie in the phonetic agreement between Malagasy and Maanyan reflexes of these consonants, and not in the fact that they have fallen together. It is true that these consonants have also fallen together with **l* and **ɣ* (as they did not in Malay), but the change **d* > *r* is not uncommon among

² Malagasy, to judge by the dialects most often cited, has no consonants before pause. If a consonant appears in a position corresponding to that of a PMP final consonant, Mérina has the vowel *a* following (and not accounted for in the comparison). The written sequence *tr* (like *dr*) stands for a single phoneme.

Malayo-Polynesian languages, and it would be difficult to calculate the probability that independent change was involved. Thus 'le fait le plus significatif' of Dahl is not as significant as it might appear to be at first glance.

Dahl makes a great point of the fact that the two languages show the same phonetic irregularities in comparison with the other languages. It seems clear that in making such assertions we must be careful to avoid contradicting the theory that phonemes change ('sounds change regularly'); otherwise our entire comparative procedure will have no scientific meaning. It is precisely by the system of regular correspondences that we can say that words of different related languages were originally the same word. How then can related languages show the same phonetic irregularity? This could only occur if some change other than phonemic change were involved, such as analogic change or borrowing.

Under type (b) there are several phenomena to be listed. (1) Some $*d$ and $*ḍ >$ Mlg. and Mny. r , other $*d$ and $*ḍ >$ Mlg. tr , Mny. d ; (2) some non-final $*γ >$ Mlg. and Mny. r , whereas some non-final $*γ >$ Mlg. and Mny. zero; (3) some non-final $*t' >$ Mlg. and Mny. s , whereas some non-final $*t' >$ Mlg. zero, Mny. h ; (4) some final $*a >$ Mlg. and Mny. a , whereas some final $*a >$ Mlg. i (written y), Mny. e .

Dahl solves (1) and (2) by concluding that the words involved are loanwords, in the private proto-language of Malagasy and Maanyan, from Borneo languages in which $*d$ and $*ḍ >$ d , such as Ngaju-Dayak and Lowangan (117), and $*γ >$ r , such as Ngaju-Dayak and Malay (59). This theory will explain why it is that some words of Indic origin (like Mlg. *trosa* 'debt', Mnj. *dosa* 'sin') show this reflex and at the same time why no $*g'$ appears as r in the two languages for $*g' >$ r in Ngaju-Dayak and Lowangan.

Dahl rejects (115 f.) Dempwolff's theory that Mlg. tr reflects the 'reduzierte Nasalverbindungen' $*nd$, $*ṇḍ$. He is quite correct in his criticism that 'Dempwolff n'avait pas le point de vue historique en linguistique'. I feel, however, that he goes too far in adding, 'Pour lui les "lois phonétiques" étaient beaucoup plus des formules de comparaison synchroniques que des essais d'expliquer les transformations phonétiques des langues à travers les siècles et les millénaires.' How then would one explain Dempwolff's attempt to characterize his proto-phonemes phonetically?

If indeed these hypotheses of borrowing could be maintained, we should have incontrovertible proof of Dahl's theory. A borrowing is an innovation, and a single shared innovation is a proof of subrelationship. But whenever Malagasy and Maanyan show different corresponding reflexes for the same reconstructed phoneme we might contrariwise conclude that different proto-phonemes are reflected. We should arrive at this via the working assumption that all corresponding words are inherited material unless there is evidence to the contrary. Here Mlg. *trosa*, Mny. *dosa* do not constitute contrary evidence by themselves. If in a comparison both languages should show $*d >$ Mlg. tr , Mny. d , we should expect them to reflect the borrowed Indic word in just the forms found. With Dahl (100) we could explain the equation Sakaláva *asara* 'summer' = Skt. *aṣāḍha* 'month including parts of June and July' on the basis that 'Skr. *dh* a pu

avoir une prononciation semblable à *r* [of the borrowing language or languages¹] en *prākṛit*'. Such a solution is as satisfactory as Dahl's. The fact that only a few etymologies involve Mlg. *tr* is not by itself convincing. The same sort of argument applies in the case of Mlg. *r*, Mny. *r* for **γ*. We are hardly prepared at this point in the reconstruction of Proto-Malayo-Polynesian to reach a hypothesis of borrowing from dialects or closely related languages where two or more languages show different corresponding reflexes. The theory that Mlg. *tr* < **d* indicates a borrowing is thus a hypothesis which is made possible by the main hypothesis of the subrelationship of Malagasy and Maanyan, but has no other relation to that hypothesis.

The latter approach to the dual reflexes appearing in Malagasy and Maanyan for certain of the PMP phonemes reconstructed by Dempwolff is applicable in the other instances cited under (b) above. Dahl goes as far as to set up *t'*₁ and *t'*₂ to deal with (3), but he could have done the same with the other instances. This approach would be necessary without reference to the question of the sub-grouping of Malagasy and Maanyan.

If, however, the dual reflexes are assigned to different proto-phonemes, their weight for the subgrouping is clearly reduced. Whatever weight is involved is to be derived from the fact (1) that the corresponding reflexes are phonetically similar and (2) that there has been a falling together of corresponding phonemes. Since there is no reason to believe otherwise, any change attested or posited in the phonetics of one language can occur in the development of another language or again in the development of the same language. Since there is no method of determining the factor of chance involved in such a recurrence, we have no way of assigning a weight to phenomena of the type described in (1). The phenomena described in (2) are different in that a phonemic distinction is lost in at least one position. We may confidently assign greater weight to the phenomena of type (2) than to those of type (1). But it is interesting to observe that if the discussion be restricted to the changes specifically referred to by Dahl in the long passage quoted, we arrive at the fact that **d, d, g' > (*d >)* Mlg. and Mny. *r* in non-final position, whereas in final position they > Mlg. *tr(a)*, Mny. *t*. We need not include here the results of some **γ* (or rather *γ₂*) which yield Mlg. and Mny. zero, and *t'*₂ which yields Mlg. zero, Mny. *h*, because the phonetic result is different in the two languages. The falling together here is restricted to Malagasy.

A great deal of discussion could be spent on the weight to be given to such instances of falling together, taken either separately or together with other factors. What is involved is of course an attempt to calculate the chance factor, and this calculation is probably too complex to be attempted.

To this calculation must also be added the fact that both languages exhibit the falling together of **b* and **w* in Mlg. *v*, Mny. *w*, and also the regressive assimilations **l-r >* Mlg. and Mny. *rr-*, and **r-l >* Mlg. and Mny. *ll-*. Of these changes, however, the first two are found in other languages at least in part (e.g. NgD. *rr* < **l-r* and NgD. *w* < intervocalic **b*).

5. Dahl mars his work by introducing the hypothesis that 'un changement

¹ See Pischel, *Grammatik der Prakrit-Sprachen* 170.

phonétique ne frappe pas tous les mots en même temps, mais saisit les mots un à un' (42, cf. 60). This theory is derived from Sommerfelt and Adriani. It is in no way different from the notion of 'sporadic change'; for if a phonetic change affects words one at a time, it can cease at any point. The net result of this theory is to make comparative studies pure guesswork, since the results of phonetic change are no longer an intertwined complex of relations between the phonemes of different languages, but simply a mess of hit-or-miss correlations intuitively reached by the investigator. In effect such a theory, if actually applied, would invalidate Dahl's entire procedure.

Fortunately it is only the formulation that is at fault here, not the comparative work. Dahl is attempting to formulate a step which will permit him to deal with the effects of 'dialect borrowing'. This can be seen in the quotation below. It is apparently his intention here to explain how PMP *t'* could have become **h* in the proto-language of Malagasy and Maanyan, and at the same time in that same language could have become **s*. He says (42 f.):

Je suppose donc qu'il s'est trouvé, au sein de la langue de laquelle dérivent le maanjan et le malgache, un centre où le changement *s > h* a commencé dans quelques mots, et d'où il s'est propagé. Comme ce ne sont pas toujours les mêmes mots qui ont changé l'*s* dans nos deux langues, je suppose qu'elles ont évolué de dialectes différents de la langue mère. Le changement phonétique a dû, à un certain degré, saisir différents mots dans ses voies de propagations différentes. Nulle part la tendance de changement n'a abouti à changer tous les *s* en *h*. Elle s'est arrêtée à un certain moment, l'*s* est demeuré un phonème dans nos deux langues. Il est impossible de dire ... pourquoi la tendance de changement n'a pas abouti. Mais l'influence des langues voisines, qui ont conservé l'*s* a pu jouer un grand rôle.

To reformulate this statement so that it will not invalidate his procedure, we need only omit the phrase 'dans quelques mots' in the first sentence. As for the beginning of the second sentence, we should rather say that Mlg. *s* corresponding with Mny. *s* reflects **t'*₁, while Mlg. zero corresponding with Mny. *h* reflects **t'*₂. Those instances which contradict this formulation have a different explanation; a resort to dialect borrowing may be necessary here. It should be kept in mind that the Malayo-Polynesian comparative formulas thus far arrived at are least satisfactory in their application to Malagasy; this only means that our hypotheses are insufficient, not that phonetic changes are sporadic.

6. I have now reviewed the main arguments presented by Dahl in support of his thesis. To strengthen his thesis he draws an analogy between the results of his investigation and those which might appear if one were to compare Icelandic with High German and Norwegian (371):

Si l'on compare les langues germaniques modernes, on trouve que le haut allemand et l'islandais sont les plus riches en formes grammaticales. Les catégories exprimées par les formatifs sont généralement identiques, et souvent les formatifs eux-mêmes.

Le norvégien a une morphologie beaucoup plus simple et, si l'on ne jugeait que d'après la morphologie, on pourrait trouver a plus d'affinité avec le haut allemand qu'avec le norvégien. Nous savons pourtant que l'islandais s'est séparé du norvégien par une émigration au 10^e siècle.

Si par contre, on compare l'évolution phonétique et le vocabulaire de ces langues, on voit que l'islandais et le norvégien ont la plus grande concor-

dance. Les formatifs que le norvégien possède encore, ont également tous leurs parallèles dans la grammaire islandaise.

Notre examen des langues indonésiennes a donné un résultat pareil. Les formatifs *maanjan*, qu'on ne retrouve pas en malgache ont pu être empruntées aux langues voisines. Par contre, le *maanjan* a des formatifs fossilisés qu'on retrouve à l'état vivant en malgache.

The notion expressed here that formatives can be borrowed does not necessarily contradict the widespread assumption that such items tend to resist borrowing. This assumption is generally applied in the proof of relationship, not necessarily of subrelationship. It is quite conceivable that the borrowing of formatives could occur rather more easily (and hence probably more frequently) between related than between unrelated languages.

Though the analogy is granted, however, the conclusion does not necessarily follow; for I believe it is easy to imagine other developments: within the procedures described, the chosen hypothesis is not the only one that will satisfy the facts. A number of phonetic changes, including some instances of the falling together of phonemes in all or some positions, have been shown to coincide with a striking number of exclusively shared items of vocabulary. But I think other comparatists would agree with me in doubting that this evidence constitutes proof. The difficulty here, as for example in the case of a proposed Italo-Celtic unity, is that the chance factor involved in each individual shared feature and in all the shared features taken together has never been calculated. Dahl has not presented a single instance of the type which must be viewed as a shared innovation; but his conclusion is not one that can be reached by 'inspection'.

7. The third procedure for subgrouping is that implied by 'lexicostatistic dating', best explained in an article by its originator Morris Swadesh, *Lexicostatistic dating of prehistoric ethnic contacts*.⁴ The procedure is essentially based on the observation that the rate of replacement of the basic vocabulary is constant. It is more convenient, however, to handle the converse, the rate of retention of basic vocabulary (i.e. inherited vocabulary), for which the constant has been determined to be $81\% \pm 2\%$ per 1000 years: at the end of a 1000-year period, a language loses about 19% of its basic vocabulary and retains about 81%. From this constant one can calculate the rate of retention on the part of two related languages developing separately. It is expected that two related languages which are separated by 1000 years from their unity will show 81% of $81\% = 66\%$ of the items in their test lists as shared inheritances. The period of differentiation can be calculated for any percentage of shared inheritances and thus the time at which any two related languages were the same language can be calculated.

Following this procedure as well as I could, but avoiding problems of detail, I arrived at the following interesting results: of its basic vocabulary *Mérina* shares 45% with *Maanyan* (implying a differentiation period of 1900 years), 28% with *Malay* (d.p. 3020 years), and 26% with *Ngaju-Dayak* (d.p. 3210 years); whereas *Maanyan* shares 39% with *Ngaju-Dayak* (d.p. 2220 years) and 36% with *Malay* (d.p. 2620 years). These figures agree with Dahl's hypothesis that the last unity

⁴ *PAPS* 96.452-63 (1952).

of Malagasy was with Maanyan. — It has been suggested that Malagasy is closely related with Sumatran Batak. But Toba-Batak shares 20 % of its basic vocabulary with Mérima (d.p. 3800 years) and 31 % with Maanyan (d.p. 2800 years).

The evidence from lexicostatistic dating, if it is worth anything, is independent of the other evidence. For this reason its results cannot be said to 'confirm' Dahl's hypothesis; rather, it appears to replace his evidence completely. In effect, the procedure of subgrouping by lexicostatistic dating is the only one available where 'inspection' or the search for a shared innovation yields no result. I had found Dahl's hypothesis wholly dubious until I submitted the languages to lexicostatistic dating. To me this means that the procedure, now in its infancy, deserves our attention and support; for it may lead toward the solution of a number of critical problems in comparative linguistics. What is needed is thorough checking against information derived from other sources in order to develop some notion of its reliability. A conceivable consequence is that Dahl's accumulation of material may after all turn out to be sufficient for his purpose.

Although the etymological bias naturally introduced into my Maanyan list by using Dahl's material was reduced by also using S. H. Ray's material,⁵ it was perhaps not completely eliminated. Thus the differentiation period of 1900 years for Malagasy and Maanyan may be too short. But the beginning of the differentiation of the two languages need not coincide with the date of the Malagasy emigration: the two languages may have begun to differentiate before that date. Thus Dahl's date of approximately 400 A.D. may be right. What he has shown conclusively, however, is only that the departure of the Malagasy must follow the Hindu penetration of the East Indies, which is generally believed to have begun in the 1st or 2nd century A.D. Whether the date should be closely associated with that of the Muara Kaman inscriptions, as it is by Dahl, is anybody's guess. As for the latest date, a great deal depends on how one chooses to view the Arabic report of the 13th century quoted above. The remainder of Dahl's discussion of this point is a collection of bits and pieces, woven together with great skill.

From the linguistic point of view this book is interesting because it attempts to handle the problem of subgrouping in more or less the traditional manner. Some may believe that Dahl proves his point; but the procedure leaves wide room for controversy because it cannot calculate the chance factor. The agreement with the results from lexicostatistic dating apparently removes all doubt, provided indeed that that procedure itself is reliable.

Anthropology today: An encyclopedic inventory. Prepared under the chairmanship of A. L. KROEBER. (International symposium on anthropology, Wenner-Gren Foundation for Anthropological Research: The inventory papers.) Pp. xv, 966. Chicago: University of Chicago Press, 1953.

Reviewed by DAVID L. OLMSTED, *Yale University*

This volume contains the prepared papers delivered at the International Symposium on Anthropology, 9-20 June 1952 in New York, sponsored by the

⁵ Languages of Borneo, *Sarawak Museum journal* 1:4.1-196 (1913).

Wenner-Gren Foundation for Anthropological Research. It consists of a preface by Paul Fejos, an introduction by A. L. Kroeber, and fifty papers, apportioned among three sections: Problems of the Historical Approach (with subdivisions: Method, Results, Theory), Problems of Process (same subdivisions), and Problems of Application (Results).

Probably few anthropologists would claim the specialized knowledge necessary for an even-handed review of all aspects of this great compendium. Certainly the reviewer makes no such claim. Moreover, the companion volume¹ may be expected to provide a definitive review in the words of the symposium's participants themselves. These considerations will justify an incomplete treatment of some papers and a disproportionate emphasis on those which are more closely linked to linguistics by subject matter or theory.

The rapid coalescence of archeology, linguistics, and physical and cultural anthropology has been chronicled by Howells, who also notes the centrifugal force of increasing specialization in these areas:²

During the last two generations at least, the rewards have gone, in any science, to the specialist, and the exploitation of the several fields of anthropology at a pace demanded by their early promise could only be carried out by men extensively trained in and devoted to these fields. Goldenweiser pointed out that in the future no single individual could be expected to comprehend all of anthropology; he designated Boas as the first true general anthropologist, and ventured the guess that Kroeber was likely to be the last.

It is fitting that Kroeber should have been chairman of the committee that planned the symposium and edited the volume of contributions. A prodigious worker for more than half of the century-long existence of anthropology as a separate discipline, he is accorded due recognition in these pages. Of the forty-five papers with reasonably complete bibliographies or citation of sources, Kroeber appears as an authority in no less than fifteen, a fact the more remarkable in that the older works of other authors are seldom referred to. In the same sample, the works of Boas are cited in eight papers, those of Kluckhohn in thirteen.

Kroeber is credited, for example, with fundamental contributions to the classification of the Australian languages (279), with being one of the first to introduce the principle of correlation of artifact change with relative depth (365), with first developing seriation or 'horizontal stratigraphy' (366), with addition of the concept of the 'horizon style' (375), with first use (in 1907) of the phonograph for the recording of interview materials, and with having been one of the few anthropologists to do research in experimental phonetics (917). One paper (that of Redfield) begins, 'The treatment derives from understanding of the subject that has been provided by Kroeber in a long series of his publications; indeed, there is no fundamental idea that I can find to add to his.'

¹ Sol Tax, Loren C. Eiseley, Irving Rouse, and C. F. Voegelin (edd.), *An appraisal of anthropology today* (Chicago, 1953). This huge volume of discussion (395 pages of text) was made available to the reviewer only after this review was substantially complete. The book contains comments by the participants in the conference; it is to be recommended for a fuller discussion of the several papers by scholars who are, unlike the reviewer, peers of the writers themselves in their various specialized fields.

² W. W. Howells, *The study of anthropology*, AA 54.3 (1952).

The papers by archeologists (Heizer, Oakley, Rouse, Movius, Childe, Bennett, Caso, Krieger, Clark, Willey) are in the reviewer's opinion easily the most uniformly excellent. Those who have been inclined to think of archeologists as mere shard-sharks will be surprised at the careful definition of terms, the explicit statement of assumptions, the relevance of theory to immediate research projects, and the clear separation of speculative from factual statements which distinguish these papers.

The contributions on physical anthropology, while somewhat uneven in quality, nonetheless reach a high level on the average; noteworthy among them is S. L. Washburn's review of the 'new' physical anthropology (*The strategy of physical anthropology*, 714-27), one of the outstanding contributions in the volume.

The majority of the papers on cultural anthropology (or anthropology as a whole) provide thorough reviews of the topics assigned; they are well written and soundly documented. In general, those that touch upon linguistic subjects (e.g. Clyde Kluckhohn's article: *Universal categories of culture*) show considerable linguistic competence. Three papers are selected for more detailed discussion here because they either deal with problems of 'communication' or raise questions that linguists face along with other students of human behavior. These are Claude Lévi-Strauss, *Social structure* (524-53); Margaret Mead, *National character* (642-67); and David Bidney, *The concept of value in modern anthropology* (682-99).

Lévi-Strauss's paper seems to the reviewer to be of uneven quality. Particularly interesting and revealing are his good-tempered discussions of other work in this field, where there have been so many vigorous disputes. But Lévi-Strauss's own formulations are disappointing. He makes much use of the term 'structure', and is at pains to define it, since so much of his argument depends on it:

The term 'social structure' has nothing to do with empirical reality but with models which are built up after it. ... a structure consists of a model meeting with several requirements. First, the structure exhibits the characteristics of a system. It is made up of several elements none of which can undergo a change without effecting changes in all the other elements. In the second place, for any given model there should be a possibility of ordering a series of transformations resulting in a group of models of the same type. In the third place, the above properties make it possible to predict how the model will react if one or more of its elements are submitted to certain modifications. And, last, the model should be constituted so as to make immediately intelligible all the observed facts (525).

Leaving aside the question of how something may be required to account for the observed facts and yet have 'nothing to do with empirical reality', the whole difficulty seems to inhere in agreement on what constitutes making the facts 'immediately intelligible'. The first three requirements that Lévi-Strauss would have his model meet are unobjectionable for a formal system like mathematics. Social 'facts', however, are made intelligible only by accounting for them in terms of theories which purport to predict (even if in limited fashion) and in this sense 'explain' them. No objection is made here to the use of models per se, but only to their introduction by a questionable analogy with mathematics, physics,

communication theory, and other disciplines rather than by some theory which relates to sociocultural data. For a legitimate use of models in behavior theory, the reader is referred to the bibliographies in Marx³ on the continuing discussion between Hull and Tolman and their followers, where every model is tested by the predictability of some behavioral event.

That his models do not arise from an explicit behavior theory is amply attested by Lévi-Strauss, who writes:

On the observational level, the main—one could almost say the only—rule is that all the facts should be carefully observed and described, without allowing any theoretical preconception to decide whether some are more important and others less (526).

The contrast could hardly be more striking between such raw empiricism and the theories summarized in this volume by Martinet and Willey. Lévi-Strauss's position is reminiscent of the 'old' physical anthropology reviewed by Washburn, where 'the facts speak for themselves'. One can only assert again that nobody has ever collected 'all the facts', or described ANY facts in the absence of some assumptions, stated or not. The point is that stating one's assumptions in advance allows one to weed out those that are inconsistent, to give operational definitions of concepts, and to derive hypotheses.

Since the notion of 'structural model' is not constrained by development from an articulated body of assumptions and hypotheses, Lévi-Strauss is free to predicate almost anything of it. For example, 'a structural model may be conscious or unconscious without this difference affecting its nature' (527). Surely the mathematical analogy has been left far behind at this point, for what could be the transformation rules for an unconscious structural model? Similarly, it is said that sometimes a model can act as a 'screen to hide' a structure which 'does not lie at a great depth' (527). Models also turn up in the 'collective consciousness', and are 'constructed by the culture' (527). Such models might 'provide some insight into the structure of the phenomena' (527). But structures were defined as models having certain properties!

It seems clear that Lévi-Strauss distrusts the observable facts; he puts his faith in 'unconscious reality' (527) and writes (534):

... among the Bororo, spacial configuration reflects not the true, unconscious social organization but a model existing consciously in the native mind, though its nature is entirely illusory and even contradictory to reality.

How can such verbalization lead to research? Rather, it leads Lévi-Strauss to an analogy between the organization of kinship and the use of language in communication. He speaks of the 'communication of women' in marriage throughout the society, of low-speed communication (women circulating in the breeding population) and high-speed communication (language). The utility of thus extending the term 'communication' is not apparent. In short, the reviewer still subscribes to the view of Kroeber (quoted 524):

'Structure' appears to be just a yielding to a word that has a perfectly good meaning but suddenly becomes fashionably attractive for a decade or

³ Melvin H. Marx (ed.), *Psychological theory* (New York, 1952).

so—like 'streamlining'—and during its vogue tends to be applied indiscriminately because of the pleasurable connotations of its sound. Of course a typical personality can be viewed as having a structure. But so can a physiology, any organism, all societies and all cultures, crystals, machines—in fact anything which is not wholly amorphous has a structure. So what 'structure' adds to the meaning of our phrase seems to be nothing, except to provoke a degree of pleasant puzzlement.

Mead's paper is a spirited defense of 'national character' studies, a field in which she has been a pioneer. She takes up criticisms of the method one by one and gives her answers forthrightly. The reviewer's reservations pertain to the great eclecticism which has characterized students of 'national character' with respect to theory. Mead writes:

Material ... can be used in combination with the theoretical schemes of personality and culture research to amplify the personality and culture conceptualizations by intensive elaboration of theory from other sources (652).

She goes on to specify some of these other sources: cultural theory, learning theory, Gestalt psychology, Freudian psychology, child development studies (651), literary criticism, architectural history, equilibrium theory, content analysis, biophysics, cybernetics, sociometry, topological psychology, 'and so forth' (653). It will be a major task to integrate these schemes, which are in different stages of formulation, with varying degrees of rigor in their derivations, sometimes conflicting fundamental assumptions, and very different kinds of problems as their goal. Picking a term here and a concept there, folding in a pinch of statistics and a dash of linguistics, some students of the subject have come up with a dish of conclusions that seem oversimplified and largely unacceptable to workers in related fields. As an instance, it seems premature to jump from such a theoretical base (still undeveloped) to an 'explanation' of the bombing of Manila after it was declared an open city, or of attempts by Puerto Ricans to assassinate the President.

Laudably, Mead states her own assumptions at some length in a not unsuccessful attempt to begin the process of genuine integration. Not all workers will agree with all or even most of these assumptions, but their clear statement is a step in the right direction.

Bidney's paper, after an introduction that is largely a history of some aspects of philosophy going back to Aristotle and the Stoics, devotes a good deal of space to a critique of cultural relativism, mainly as presented by Herskovits. The reviewer has no wish to get into this argument, but he must take exception to Bidney's view of the anthropologist's duty (687): 'How to find substitutes for traditional religion, which will promote the feeling of solidarity and peace of mind which religion formerly produced, remains an unsolved ethnological problem.' This may be somebody's problem, but it is not the ethnologist's. Again, Bidney speaks (689) of 'the naive optimism of cultural laissez faire' and 'the need for cultural integration on a world scale'. As Bidney sees it (694), 'The only effective alternative to a mythical relative absolute is a better, more rational, and more objective ideal of conduct and belief, capable of overcoming the limitations of the former.' This may or may not be the case, but is it the task of

the anthropologist to determine what is 'better', 'more rational', and 'more objective'? The reviewer would certainly shrink from such a task, except as a private citizen whose vote carries no more weight than the next man's. The point is admirably made by G. Wilson, the first director of the Rhodes-Livingston Institute (quoted by Forde 858):

... the social anthropologist is entitled, as a man, to his own moral and political views—they are no more and no less worthy of respect than those of any other well informed citizen—but he is not entitled to pass them off as 'scientific'. ...

Human societies, like the earth on which they live, have a hard material reality which cannot be mastered without patient and objective study. It is the scientists' business to undertake that patient and objective study, it is the business of government and industry to make use of their results in fashioning out of the present whatever future they desire. The scientists must make it their boast that both governments and oppositions can trust them equally because they say nothing that they cannot prove, because they are always pedestrian and never leave the facts. The men of affairs must make it their boast that they allow the scientists perfect freedom in their researches and pay to their results when published the attention which proven fact deserves.

The papers on linguistics which deal with basic research as opposed to application are Joseph H. Greenberg, *Historical linguistics and unwritten languages* (265-86); Floyd G. Lounsbury, *Field methods and techniques in linguistics* (401-16); Harry Hoijer, *The relation of language to culture* (554-73); and André Martinet, *Structural linguistics* (574-86). Hoijer's paper, after a general consideration of the topic, summarizes the Whorf position and various hypotheses derived from it. It need not be discussed here, since a group of linguists and scholars from associated disciplines, meeting at the University of Chicago in March and April 1953, dealt extensively with this very topic, and since their results are to be published as a monograph of the American Anthropological Association.⁴

Greenberg presents, in detail, the assumptions generally used by linguists in the application of historical linguistics to unwritten languages. He reviews the controversy between Meillet, Boas, and Schuchardt, and presents the views of Trubetzkoy. Greenberg speaks of 'a rigorous application of Sapir's insight' (284); much of the paper is devoted to making explicit and objective, by means of probability hypotheses, just such insights as Sapir is famous for. In this sense the paper is a very considerable step forward, and represents an original contribution rather than a summary. Greenberg applies the techniques to selected regions: he presents his well-known sound and increasingly accepted classification of the languages of Africa; reviews the Australian evidence and previous classifications, concluding that all the languages of that continent form a single family; presents his views—the result of a 'cursory investigation'—on America north of Mexico; and outlines a new classification for Southeast Asia.

Greenberg rejects Schmidt's Austric hypothesis and lists the families of South-

⁴ The participants were Hoijer (chn.), Greenberg, Hockett, Lounsbury, McQuown, Lenneberg, Fearing, Kennard, Roberts, Kaplan, Newman, Voegelin, Singer, Kroeber, Eggan, Redfield, von Gruenebaum, and Wright.

east Asia as follows: (1) Sino-Tibetan, (2) Austro-Asiatic, (3) Kadai-Malayo-Polynesian, (4) Andaman Islands, (5) perhaps Nehari. He places Thai with Kadai-Malayo-Polynesian, and assigns both the Miao-Yao dialects and Min-Hsia to Sino-Tibetan. Rejecting Maspero's contention that Annamite has a close relationship with Thai, Greenberg puts it firmly in the Austro-Asiatic family.

Lounsbury's paper is a practical discussion, well suited to the general anthropological reader. Yet it manages to include an outline of the main principles of descriptive analysis, as well as advice on the strategy of working with an informant and other 'field' techniques. Few linguists will quarrel with his conclusion (403) that 'skill in phonetic transcription ... should be regarded as the minimum essential in the linguistic training of the field anthropologist'.

Lounsbury provides a good explanation of what the linguist actually does. His use of terms like 'sound-type' enables him to make a good case for phonemic organization as a field technique as well as an analytical tool to be used later. This argument is so carefully worded that the general reader hardly realizes he is applauding his old bugaboo the phoneme. Recognizing that field techniques have to be chosen with regard for their relevance to particular problems, Lounsbury tells the nonspecialist how to get data for glottochronology, the measuring of dialect distance, dialect geography, semantic studies, and other topics that may interest the ethnographer or the culture historian.

Martinet's paper is a dispassionate and enlightening presentation of the main points of agreement and disagreement between Trubetzkoyan, Hjelmslevian, and Bloomfieldian linguists. He devotes some effort in attempting to define 'structural' linguistics; and though he does not quite succeed in bringing it off, he is yet able to show clearly where the different 'schools' agree and where they differ. The reviewer is inclined to wonder whether Kroeber's dictum on 'structure' is not applicable here also. For example, a clear definition of 'structural' would clear up the reviewer's difficulties with this sentence: 'It is a fact, however, that the Bloomfieldians' main concern with analysis contrasts with the constant emphasis placed by other schools on the structural nature of language'. The 'structuralism' of such a member of the Prague School as Trnka⁶ does not contrast with analysis, but rather, presupposing it, purports to explain changes through time, very much like Malinowski's functionalism.

Martinet adopts the refreshing practice of explaining Trubetzkoyan and Bloomfieldian notions and practices in Hjelmslevian terms, instead of following either of the other, more familiar paths. Adherents of the various 'schools' may disagree with points of emphasis, but in the main, Martinet's review seems characterized by Olympian detachment. One criticism, admittedly carping: Martinet, perplexed with the problem of what to call followers of Bloomfield, tries 'Bloomfieldian' and 'American', but finally settles on 'the Yale School'. This is surely an unfortunate choice, since Sapir, whom Martinet puts on the side of the 'structuralists', spent more active years and trained many more students at Yale than Bloomfield, whose approach Martinet calls 'more analytical'. In view of the way in which Bloomfield's influence was exerted, namely through his writings,

⁶ Bohumil Trnka, *Hláskoslovné zákony v strukturálním jazykozpytu*, *Časopis pro moderní filologii* 23.385-8 (1937).

the term 'Bloomfieldian', awkward as it is, would seem more apt than any other, parallel to 'Saussurean' rather than 'Genevan'.

Of the papers on problems of applications, the one by Mary R. Haas (The application of linguistics to language teaching, 807-18) will be of most interest to readers of this journal. She outlines the new emphases in language teaching which have come increasingly to the fore in the last decade; traces their development from the research of Boas, Sapir, and Bloomfield and their wider acceptance during the late war; and reports on their present reputation among the general company of language teachers.

The other papers on applied anthropology, together with Rowe's valuable contribution on technical aids, give adequate coverage to their topics. A few subjects that failed to receive any treatment are the use of linguistics as the basis of new orthographies and related problems (R. A. Hall's trip to Haiti for UNESCO, for example), pipe-line archeology on the Navaho reservation, Mead's experiments with photography and kinesic interpretation, and the use of linguistics in aphasia studies, mentioned only in passing.

A footnote, unaccountably appearing on page 488, explains that the plan of bibliography was changed too late for all contributors to amplify their lists and refer to each item specifically in the text. The result is that the bibliographies are of uneven quality.

At their best, the contributions in this volume are superb guides to the literature, and they maintain throughout a remarkably high standard of quality. It is not too much to say that these papers comprise the most complete and the most authoritative corpus of remarks on anthropology ever put between two covers. Both what is peculiar to the several subdivisions and what is shared are much in evidence. The methods and assumptions required by the several bodies of phenomena to be investigated are distinguished from the preoccupation with theory-construction, validation, and duplicability of results that is common to students of all branches of anthropology. Hockett⁶ has remarked that every linguist is an anthropologist whether he specifically designates himself one or not; this is an almost perfect book⁷ for the linguist who would like to learn the nature of the charge before deciding how to plead.

⁶ AA 52.113 (1950).

⁷ The printing and manufacture of the volume are excellent, particularly in view of the admirably short time that elapsed between the conference and the publication of the book. The following errors were noted; references are to pages (a and b refer to the first and second columns):

213a line 17 *distributions* for *distribution*; 222a line 3 from bottom *ceramc* for *ceramic*; 265 fn. *Hjelmaler* for *Hjelmslev*; 266 fn. *Twaddel* for *Twaddell*; 268a line 31 *French or Italian* for *French and Italian*; 271b line 22 *languages* for *language*; 274a line 1 *Hindustania* for *Hindustani*; 286a line 31 *Algonkin* for *Algonquian*; 325 b line 3 from bottom *Bernard* for *Bernhard*; 326a line 1 *Bernard* for *Bernhard*; 605b line 8 *Frenkel*, for *Frenkel*; 635b line 14 *Edward Spier* for *Edward Spicer* (reviewer's best guess; some articles referred to were in the Edward Sapir volume, of which one of the editors was Leslie Spier; the bibliography lists nothing by Sapir, and only one paper by Spier, dated 1929, whereas the article in question must be either 1941 or 1943; Spicer's paper, listed in the bibliography, is not referred to elsewhere in the chapter, and is dated 1943); 751a line 2 *wo* for *who*; 754b line 23 *wills* for *will*; 904a line 3 from bottom *has* for *had*.

PUBLICATIONS RECEIVED

This listing acknowledges the receipt of recent works that appear to bear on the scientific study of language. No book can be returned to the publisher, nor can the Editor promise that every book received will be reviewed in the journal. Reviews are published as circumstances permit, and copies are sent to the publishers of the works reviewed.

- Acta linguistica Academiae Scientiarum Hungaricae 2.247-496 (1953).
 Aegyptus 32.241-496 (1952).
 Afrika und Übersee 37.49-144 (1953).
 Anales del Instituto nacional de antropología e historia 4.1-322 (1949-50, publ. 1952); 5.1-195 (1951, publ. 1952).
 Ankara Üniversitesi: Dil ve Tarih-Coğrafya Fakültesi Dergisi [Université d'Ankara: Revue de la Faculté de Langues, d'Histoire et de Géographie] 7.255-709, 1-97 (1949); 8.1-686 (1950); 9.1-478 (1951).
 Annals of the Bhandarkar Oriental Research Institute 33.1-316 (1952, issued 1953).
 A nyelvtudományi intézet közleményei [Communications de l'Institut linguistique de l'Académie des Sciences de Hongrie] 2.4.1-193 (1951), 4.1.1-222 (1953). Budapest: Akadémiai Kiadó.
 Die Apokope des mhd. -e in seinen verschiedenen Funktionen. By Kaj B. Lindgren. (Annales Academiae Scientiarum Fennicae, Ser. B, Vol. 78.2.) Pp. 225. Helsinki, 1953.
 Archiv orientální 20.1-671 (1952) = Diatribae quas amici collegae discipuli Francisco Lexa quinque et septuaginta annos nato ddd; ed. Zbyněk Žába.
 Bahasa dan budaya 1:1.1-32 (1952); 1.4.1-39, Suppl. 1-32 (1953): Madjalah populer jang berdasarkan ilmu pengetahuan tentang bahasa dan budaya. Lembaga bahasa dan budaya, Fakultas Sastra dan Filsafat, Universitas Indonesia di Djakarta.
 Beiträge zur Erforschung des Luvischen. By Bernhard Rosenkranz. Pp. [viii], 100. Wiesbaden: Otto Harrassowitz, 1952.
 Bibliographie linguistique de l'année 1950 et complément des années précédentes. Publié par le Comité International Permanent de Linguistes ... sous les auspices du Conseil International de la Philosophie et des Sciences humaines. Pp. xxviii, 275. Utrecht-Anvers: Spectrum, 1952.
 Bibliography of Slavic folk literature. By William E. Harkins. (Columbia Slavic studies.) Pp. vii, 28. New York: King's Crown Press, Columbia University, 1953.
 Biblos 28.1-596 (1952).
 Bulletin analytique: Philosophie 7.1-450 (1953).
 Bulletin of the Department of Anthropology 1.1-159, plates i-xviii (1952). Government of India, Calcutta.
 Bulletin of the Ramakrishna Mission Institute of Culture 4.55-102 (1953). Calcutta.
 Bulletin of the School of Oriental and African Studies 15.203-422 (1953).
 The category of person in language. By Paul Forchheimer [with a preface by G. Bonfante]. Pp. [xiv], 142. Berlin: Walter de Gruyter and Co., 1953.
 The classical weekly 46.161-232 (1953).
 Cultura neolatina 12.81-184 (1952).
 Einführung in die altaische Sprachwissenschaft; 2. Formenlehre. By G. J. Ramstedt, ed. Pentti Aalto. (Mémoires de la Société finno-ougrienne, No. 104.2) Pp. 262. Helsinki: Suomalais-Ugrilainen Seura, 1952.
 English studies 34.49-96 (1953).
 Escorial Bible I.J.4. Ed. by O. H. Hauptmann. Vol. 1, The Pentateuch, pp. xii, 320. Philadelphia: University of Pennsylvania Press (for Grinnell College Press), 1953.
 Étude de phonétique auditive sur les parlers de l'Algarve. By Göran Hammarström. (Diss. Uppsala.) Pp. 187. Uppsala, Stockholm: Almqvist & Wiksells Boktryckeri, 1953.

Études germaniques 8.97-240 (1953).

Gesprochenes Deutsch, 35th ed. By Wolfhart Klee and Magda Gerken. (Goethe-Institut, München.) Pp. xvi, 223. Bremen: Extraneus-Verlag (August-Wilhelm Dohse), 1953.

Gramática rarámuri. By David Brambila, S.J. Pp. xxix, 645. México, D.F.: Buena Prensa, 1953.

Grammaire de la langue védique. By Louis Renou. (Collection 'Les Langues du Monde'; ouvrage publié avec le concours du Centre National de la Recherche Scientifique.) Pp. 454. Lyon and Paris: IAC, 1952.

Haitian Creole: Grammar, texts, vocabulary. By Robert A. Hall Jr. (The American anthropologist, Vol. 55, No. 2, Part 2 = Memoir No. 74 of the American Anthropological Association, Memoir No. 43 of the American Folklore Society.) Pp. 309. Menasha, Wis., 1953.

Harvard journal of Asiatic studies 15.285-530 (1952); 16.1-292 (1953).

Hespérus 38.265-514 (1951); 39.1-574 (1952).

Hethitisches Wörterbuch. By Johannes Friedrich. (Indogermanische Bibliothek, 2. Reihe.) Fasc. 2, pp. 97-192. Heidelberg: Carl Winter, Universitätsverlag, 1953.

Histoire de la langue française des origines à nos jours. [By Ferdinand Brunot.] Tome 13: L'époque réaliste; Première partie: Fin du Romantisme et Parnasse. By Charles Bruneau. Pp. xvi, 384. Paris: Librairie Armand Colin, 1953.

Historische Lautlehre des Lateinischen. By Max Niedermann. (Sprachwissenschaftliche Studienbücher.) 3d ed. revised. Pp. vii, 214. Heidelberg: Carl Winter, Universitätsverlag, 1953.

Homenaje a Fritz Krüger. [Ed. by Toribio M. Lucero and Alfredo Dornheim.] (Universidad Nacional de Cuyo, Facultad de Filosofía y Letras.) Vol. 1, pp. xxx, 466, with a portrait. Mendoza, República Argentina, 1953.

Indian tribes of North America. By Harold E. Driver, John M. Cooper, Paul Kirchhoff, Dorothy Rainier Libby, William C. Massey, and Leslie Spier. (Indiana University publications in anthropology and linguistics, Memoir 9 of the International journal of American linguistics; Supplement to IJAL, Vol. 19, No. 3.) Pp. [iii], 30, with two folding maps in pocket. Baltimore: Indiana University, 1953.

International anthropological and linguistic review 1.1-106 (1953).

International journal of American linguistics 19.85-251 (1953).

Italica 30.65-128 (1953).

Izvestija Akademii Nauk SSSR 11.281-392 (1952); 12.1-296 (1953).

Jezik 1.97-160 (1953).

Jornal de filologia 1.1-88 (1953). Faculdade de Filosofia, Ciências e Letras, São Paulo.

Journal de la Société des Américanistes NS 41.221-624, plates i-cii (1952).

Journal of the Oriental Institute 2.105-302, 41-8 (1952). Baroda.

The journal of the Polynesian Society 61.177-328 (1952).

Languages in contact: Findings and problems. By Uriel Weinreich, with a preface by André Martinet. (Publications of the Linguistic Circle of New York, No. 1.) Pp. xii, 148. New York, 1953.

La langue allemande en France: Pénétration et diffusion des origines à nos jours; II. De 1830 à nos jours. By Paul Levy. (Bibliothèque de la Société des Études Germaniques, Vol. 8.) Pp. 275. Lyon and Paris: IAC, 1952.

Les langues du monde, nouvelle édition. Par un groupe de linguistes sous la direction de A. Meillet et Marcel Cohen. (Société de linguistique de Paris.) Pp. xlii, 1296, with 26 loose maps in pocket. Paris: Centre National de la Recherche Scientifique, 1952.

Lateinisches etymologisches Wörterbuch. By Alois Walde, 3d ed. by J. B. Hofmann. (Indogermanische Bibliothek, 2. Reihe.) Fasc. 20, pp. 673-752. Heidelberg: Carl Winter, Universitätsverlag, 1953.

Lengua española. By Luis Florez. (Publicaciones del Instituto Caro y Cuervo, Series minor, No. 3.) Pp. 301. Bogotá, Imprenta Nacional, 1953.

Leuvense bijdragen 42.93-136, Bijblad 57-110 (1952).

Levende talen, Nos. 169, 170 (1953).

Lingua nostra 14.1-60 (1953).

- Lingua posnaniensis* 4.1-347 (1953).
- La linguistique. By Jean Perrot. ('Que sais-je?' Le point des connaissances actuelles, No. 570.) Pp. 136. Paris: Presses Universitaires de France, 1953.
- Le livre des deux mille phrases: I. La méthode des dictionnaires de phrases; II. Questionnaire de deux mille phrases selon le parler d'un Pariesien. By Henri Frei. (Société de Publications romanes et françaises, No. 40.) Pp. 92. Genève: Librairie Droz, 1953.
- Logic and language (second series). Ed. by A. G. N. Flew. Pp. vii, 242. New York: Philosophical Library, 1953.
- Lulelappsk ordbok (Lulelappisches Wörterbuch). By Harald Grundström. (Skrifter utgivna genom Landsmåls- och Folkminnesarkivet i Uppsala, Ser. C.1.) Fasc. 11: Register, Efterskrift, Förklaringar, Rättelser, pp. 1589-711. Uppsala: A.-B. Lundequistska Bokhandeln; København: Einar Munksgaard, 1953.
- Lydverket i Solumsmålet. By Halvor Dalene. (Skrifter frå Norsk Målførearkiv, No. 3.) Pp. x, 95. Oslo, 1953.
- Le maître phonétique III.31.1-26 (1953).
- Man 53.1-80 (1953).
- Neophilologus 37.65-192 (1953).
- Norsk tidsskrift for sprogvidenskap 15.1-486 (1949).
- The Norwegian language in America: A study in bilingual behavior. By Einar Haugen. (Publications of the American Institute, University of Oslo, in coöperation with the Department of American Civilization, University of Pennsylvania.) Vol. 1, The bilingual community, pp. xiv, 317; Vol. 2, The American dialects of Norwegian, pp. vii, [377]. Philadelphia: University of Pennsylvania Press, 1953.
- Nueva revista de filología hispánica 6.209-424 (1952).
- Oneida verb morphology. By Floyd G. Lounsbury. (Yale University publications in anthropology, No. 48.) Pp. 111, with folding table. New Haven: Yale University Press, 1953.
- Orbis 2.1-290 (1953).
- Orientalia NS 22.129-328, 1*-56*, plates xxix-lxxii (1953).
- Orientalistische Literaturzeitung 48.1-336 (1953). Akademie-Verlag, Berlin.
- Palaeologia 1.179-284 (1952), 2.1-78 (1953). The Palaeological Association of Japan, Osaka.
- Problèmes de la langue minoenne. By Vladimir Georgiev. (Bolgarskaja Akademija Nauk: Otdelenie jazykoznanija, etnografii i literatury.) Pp. 196. Sofia: Izdanie Bolgarskoj Akademii Nauk, 1953.
- Results of the conference of anthropologists and linguists. By Claude Lévi-Strauss, Roman Jakobson, C. F. Voegelin, and Thomas A. Sebeok. (Indiana publications in anthropology and linguistics; Memoir 8 of the International journal of American linguistics; supplement of IJAL, Vol. 19, No. 2.) Pp. vi, 67. Baltimore: Indiana University, 1953.
- Revista de filología española 36.209-436 (1952).
- Revista de folklore [NS] 1.1-258 (1952). Organo del Instituto Colombiano de Antropología, segunda época; Bogotá.
- Revista portuguesa de filología 5.1-552, i-xiii (1952).
- Revue des études roumaines 1.1-241 (1953). Institut universitaire roumain Charles I^{er}, Paris.
- Sardische Studien: Das mediterrane Substrat des Sardischen, seine Beziehungen zum Berberischen und Baskischen sowie zum eurafrikanischen und hispanokaukasischen Substrat der romanischen Sprachen. By Johannes Hubschmid. (Romanica helvetica, Vol. 41.) Pp. 137. Bern: A. Francke AG. Verlag, 1953.
- Sborník Prací filosofické fakulty Brněnské University 1.1-194 (1952).
- The Scottish national dictionary. [Ed. by David D. Murison.] Vol. 4, Pt. 1, pp. 112. Edinburgh: Scottish National Dictionary Association Ltd., 1953.
- Shakespeare's pronunciation. By Helge Kökeritz. Pp. xv, 516. New Haven: Yale University Press, 1953.
- Smithsonian Institution, Bureau of American Ethnology, Bulletin 151: Anthropological papers, Nos. 33-42. Pp. ix, 507, with 37 plates. Washington: U. S. Government Printing Office, 1953.

- Smithsonian Institution, Institute of Social Anthropology, Publ. No. 15: Indian tribes of northern Mato Grosso, Brazil; by Kalervo Oberg (with appendix: Anthropometry of the Umotina, Nambicuará, and Iranxe ..., by Marshall T. Newman). Pp. viii, 144, with 10 plates. Washington: U.S. Government Printing office, 1953.
- Speculum 28.245-644 (1953).
- Studien zur Entwicklungsgeschichte des Frankoprovenzalischen. By Helmut Stimm. (Akademie der Wissenschaften und der Literatur [Mainz]: Abhandlungen der geistes- und sozialwissenschaftlichen Klasse, Jahrgang 1952, Nr. 6.) Pp. 160. Wiesbaden: Verlag der Akademie der Wissenschaften und der Literatur in Mainz (bei Franz Steiner Verlag), 1953.
- Studier i modern språkvetenskap 18.1-171 (1953).
- Studies in the language of Homer. By G. P. Shipp. (Cambridge classical studies: Transactions of the Cambridge Philological Society, Vol. 8) Pp. x, 155. Cambridge University Press, 1953.
- Subjunctive and optative: Their origin as futures. By E. Adelaide Hahn. (Philological monographs, No. 16.) Pp. xviii, 157. New York: American Philological Association, 1953.
- Sweet's Anglo-Saxon primer; 8th edition. Revised by Norman Davis. Pp. vii, 129. Oxford: At the Clarendon Press, 1953.
- Taal en functionaliteit: Een historisch-critische studie over de opvattingen aangaande de functies der taal vanaf de prae-humanistische philologie van Orleans tot de rationalistische linguïstiek van Bopp. By P. A. Verburg. Pp. xvi, 490. Wageningen: H. Veenman and Zonen, 1952.
- Tijdschrift voor nederlandse taal- en letterkunde 71.1-160 (1953).
- Transactions of the Philological Society 1952. Pp. 145, xii. Oxford: Basil Blackwell, 1953.
- Über das niederländische Adverbialpronomen *er*. By Gunnar Bech. (Travaux du Cercle linguistique de Copenhague, Vol. 8.) Pp. 32. Copenhagen: Nordisk Sprog- og Kulturforlag; Amsterdam: Taalinstituut de Natuurmethode, 1952.
- Verzamelde opstellen: Als feestgave aan de schrijver aangeboden bij zijn 65ste verjaardag. By G. G. Kloeke. Pp. [vii], 236, with portrait and 4 folding maps. Assen: Van Gorcum & Comp. N.V., 1952.
- Virittäjä 57.1-240 (1953).
- Vocabulario dei dialetti della Svizzera Italiana. Ed. by Silvio Sganzini. (A cura della Repubblica e Cantone del Ticino con l'aiuto della Confederazione Svizzera e con un contributo del Cantone dei Grigioni.) Fascicle 1, *a-agnesa*, pp. xxxvi, 40, with 4 maps in flap. Lugano, 1952.
- Volkskundliche und religiöse Begriffe im nördlichen Waldland von Kamerun. By Johannes Ittmann. (Afrika und Übersee, Beiheft 26 [Folge der Beihefte zur Zeitschrift für Eingeborenen-Sprachen].) Pp. [iv], 68. Berlin: Verlag von Dietrich Reimer, 1953.
- Words and sounds in English and French. By John Orr. (Modern language studies.) Pp. viii, 279. Oxford: Basil Blackwell, 1953.
- Zeitschrift für Phonetik und allgemeine Sprachwissenschaft 6.269-380 (1952).

31 August 1953